

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



THE DAWN OF TECHNOLOGY

Stone tools from Lake Turkana are earliest yet at 3.3 million years old

PAGES 294 & 310

QUANTUM PHYSICS

WAVEFUNCTION REALITY CHECK

Mathematical tools — or are they for real?

PAGE 278

BIOTECHNOLOGY

TROUBLE BREWING

Regulation needed to contain opiate-producing yeasts

PAGE 281

CRYSTALLOGRAPHY

MAPPING OUT DISORDER

The next challenge for structure determination

PAGE 302

© NATUREASIA.COM

21 Mar 2015

Vol. 525, No. 7552

THIS WEEK

EDITORIALS

VIOLENCE The necessary search for the brain's kill switch **p.260**

WORLD VIEW Open up the secret bioinformatics boxes **p.261**



HOT FISH Deep-water opah flaps fins to keep warm **p.262**

Publish or perish

Universities should release reports to show what they are doing to tackle misconduct — and funders should help them to do so effectively.

As every politician knows, it is important to address problems, but even more important to demonstrate that you are addressing them. When it comes to research misconduct, UK universities are failing on both points. To fail on the first is understandable: eradicating misconduct is difficult. It demands cultural change, education and a system of checks and balances. But to fail on the second is unacceptable, especially given that it is relatively easy to achieve.

The United Kingdom has no regulatory body to deal with research misconduct. Instead, since 2013 universities have had to adhere to a set of guidelines in order to receive grants from major funders. Called *The Concordat to Support Research Integrity*, the guidelines detail good practice and aim to strengthen the mechanisms available for investigating misconduct. They also call on universities to publish annual summaries of their formal misconduct investigations.

As we report on page 271, a survey reveals that most universities are not bothering to do so. And when they do, some of the reports are not very enlightening. One did not include the number of cases investigated, and another could not be accessed without a login. Four reports claimed that the universities had carried out zero investigations that year — an unlikely figure for any research-intensive university that takes the issue of misconduct and integrity seriously.

British universities are notoriously image-conscious, especially since the 1998 introduction of tuition fees established a marketplace, and it is understandable that many are reluctant to publish the figures. The few that do publish reports risk being singled out as having a problem, when in fact the reverse is true — such investigations show that the institution has processes to detect and deal with misconduct. But almost 2% of researchers admit to having fabricated, falsified or modified data at least once, according to a metastudy by social scientist Daniele Fanelli of the University of Edinburgh, UK (D. Fanelli *PLoS ONE* 4, e5738; 2009). Pretending that misconduct does not happen is no longer an option.

Discussion at a research-integrity conference in London last week suggested that many institutions have just been slow to publish details of their misconduct investigations, rather than aiming to avoid it entirely. It also emerged that staff who oversee research integrity in universities, and who are still working out how to ensure that their institutions adhere to the concordat, feel under-resourced.

For those universities that do have adequate systems to report and deal with misconduct, making investigation summaries public would be an easy win. Those institutions that have yet to make such systems a priority should remember that the concordat was introduced because UK systems for dealing with issues of research integrity had been judged inadequate by a parliamentary enquiry. Unlike in the United States, where the Office of Research Integrity oversees formal misconduct investigations related to research funded by the US National Institutes of Health, or Ireland, which plans to subject labs to spot-checks from auditors, UK universities have been allowed to police themselves.

When the concordat was introduced, many feared that it lacked

teeth. That many universities have so far been willing to skip around its recommendations does nothing to ease those fears. Currently, the only checks and balances are universities' statements to funders, saying how they are taking action.

Although universities are best placed to investigate and censure misconduct by their own researchers, funders can do more to help them. First, Research Councils UK and the Higher Education Funding Council for England, which have responsibility for ensuring compliance to the concordat on behalf of funders, should clarify the document's language and intentions. At present, the concordat says that universities "should" publish investigation reports. Many institutions seem to have read this as a suggestion rather than as a mandate. Funders should make clear who it is aimed at, and how they expect universities to comply.

Second, the funders should consider changing how misconduct investigations are published. Putting them on university websites that must be trawled manually and individually for figures is not ideal, either for the institutions themselves or for those who want to find the data. As well as making clear that universities must report the figures, the funders should collate and publish the reports.

Research misconduct is a fact, and institutions should not feel that they will be penalized for investigating cases promptly and fairly. The best way to change perceptions is to ensure full compliance. If every university acknowledges the issue, then the risk of being an outlier disappears, and only those institutions that choose not to publish will be the subject of suspicion and public scrutiny. ■

"Pretending that research misconduct does not happen is no longer an option."

A patent problem

Making lawsuits more risky for patent trolls is just one way to stop abuse of the system.

Earlier this year, a group of 51 legal scholars and economists sent a letter to the US Congress urging it to take action on the increasing toll of frivolous patent lawsuits. Over the past five years, they said, researchers have published more than two dozen studies on the economic consequences of patent litigation. The view that has emerged is grim: the lawsuits are hindering research and development, and slowing the launch of firms.

Less than a month later, another 40 scholars rebuffed the claims, saying that the impact of the lawsuits has been exaggerated. Furthermore, they argued, patent litigation is on the wane, and legislation to rein it

in could damage the US “engine of innovation” by weakening patent protections for inventors.

Such are the muddled waters that Congress has been navigating as it seeks to respond to the cries of technology companies and of President Barack Obama’s administration, which want to crack down on lawsuits launched by ‘patent trolls.’ No fairy tale, these entities are essentially holding firms to ransom, threatening organizations that are making use of the innovations with expensive, time-consuming lawsuits if they do not pay to license the patent. A 2013 attempt to curb such legislation met with failure last year. Lawmakers now seem to be making progress (see page 270).

Much of the scholarly debate boils down to a difficulty that has also plagued Congress: how to define a troll. Universities, too, license their patents, often for a fee, to those who want to use their researchers’ inventions to create a product or service. As such, they are considered ‘non-practising entities,’ a more-polite term than troll, but the two labels are often used interchangeably.

Scholars generally argue that universities should be considered differently because they work towards a social good and their patenting efforts spur innovation based on academic discoveries. This is in stark contrast to a troll, which accumulates weak, broad patents with the sole intent of using them to push firms into settling a lawsuit before the expense of the litigation damages their business. Lawmakers in the US Senate seem to agree with this distinction, and last month created a carve-out that excludes universities from some of the proposed measures for cracking down on patent trolls.

But the distinction has fuzzy boundaries: some universities are

highly aggressive in monetizing their patents, even licensing them to companies that are considered to be trolls (see *Nature* **501**, 471–472; 2013). Earlier this year, the Association of American Universities and the Association of Public and Land-grant Universities took a step in the right direction by urging their members not to align with trolls. Universities should heed that guidance or risk losing the faith of Congress and the

“It is important not to see patent-troll legislation as a panacea.”

public. The Senate loophole for institutions of higher education was a political necessity in the face of heavy lobbying by universities, but that lobbying would have been much less persuasive had it not been tied to widespread public trust.

As Congress has wrestled with definitions, its overall approach for deterring frivolous lawsuits has remained fairly constant: make them more risky for the plaintiff. It is a welcome change to a system that is much too easy to exploit, but it is a blunt tool that could jeopardize the ability of small firms to defend their intellectual property. And even if it succeeds in Congress, it will not tackle the underlying problem: the US Patent and Trademark Office is granting far too many vague and redundant patents. This is a particular problem for software, but affects other fields, too.

Measures to raise the bar — including a process that allows parties to challenge a patent without needing to resort to litigation — may be having an effect: the number of patent lawsuits dropped by 18% between 2013 and 2014. But it is important not to see patent-troll legislation as a panacea. Fundamental changes at the patent office remain the key to curbing abuse. ■

The kill switch

Brain researchers and social scientists are well placed to find out what makes humans murder.

Groups of humans have always slaughtered those who belong to other groups. The twentieth century was shot through with numerous examples, from the genocides of Armenians in Ottoman Turkey and of Jews in Nazi Europe to the massacres of ethnic rivals in civil wars in Rwanda and Bosnia during the 1990s. Today, the fundamentalist group ISIS is spooking the world with its willingness to butcher others who do not adhere to its extremist form of Islam.

Attempts to understand such events tend to focus on political reasons. But a conference in Paris last month dared to ask a different question: how, biologically speaking, do normally non-violent and psychologically stable people overcome the instinctive human aversion to killing when faced with circumstances of war or extremism? What drives them to participate in acts of genocide? This is arguably the biggest challenge for interdisciplinary dialogue across the fields that consider brain and behaviour.

All human behaviours originate in the brain, which computes cognitive and emotional information to decide what to do. So what, precisely, happens in that organ at the moment that a person’s natural abhorrence of harming others is computed out of the equation?

The organizers of last month’s conference at the Paris Institute of Advanced Studies — ‘The Brains that Pull the Triggers’ — deserve credit for even posing this question. It goes against another human instinct: to consider evil in moral rather than biological terms, as if identifying a biological signature in the brain might somehow be exploited as an excuse to absolve a person of his or her responsibility.

Neuroscientists have studied the abnormal condition of psychopathy in addition to components of normal cognition — such as the recognition of emotions in the faces of others — that may have a bearing on

the problem. And psychologists and sociologists have looked at the behaviour of ordinary individuals who identify themselves with particular groups and align their behaviour with that group.

The conference brought researchers from these disciplines together, along with historians who presented sobering data on the behaviour of soldiers in wartime. One presentation included documentation from post-Second World War interrogations of hundreds of untrained German reservists who were recruited to active service in 1942 and went on to slaughter tens of thousands of Jews in Poland. Transcripts revealed that their distraught commander had allowed anyone to opt out of killing — but only 1 in 10 did so.

This is tricky terrain for academics, and many researchers at the conference admitted some discomfort at being asked to consider their findings as being relevant to the neuroscience of repetitive killings. For some of the sociologists, it felt like an attempt to medicalize a social issue. For some neuroscientists, it felt like over-extrapolation of results from much simpler experiments. In the air was an uneasy feeling that such interpretations could seem superficial and trite, and could trivialize crimes against humanity.

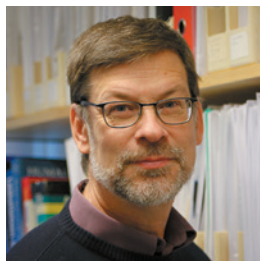
In fact, the researchers present made a brave contribution to what was a bold and important attempt to bring a multidisciplinary approach to one of the biggest questions facing humanity.

The answer will not come quickly, but research has already identified some useful paths to follow. Neurosurgeon Itzhak Fried from the University of California, Los Angeles, for example, proposes that ordinary people are able to become repetitive killers because changes in neural circuitry free the ideology-fed, cognitive parts of the brain from the emotional parts of the brain, which normally keep actions in check.

A better understanding of brain circuitry could not, of course, influence the political forces that create the conditions for mass murder. But discussion of such politically neutral basic neuroscience could allow progress while avoiding unhelpful rhetoric.

And findings in basic science could have a direct impact: perhaps by helping to find ways of educating people to make them less likely to succumb to ideological requests or commands to kill. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunqv



No more hidden solutions in bioinformatics

Precision medicine cannot advance without full disclosure of how commercial genome sequencing and interpretation software works, says Mauno Vihinen.

Last month, California became the latest in a series of places to launch a 'precision medicine' project, aiming to develop diagnostic tools and treatments based on individual genomic data. Advances in sequencing technology have already made the US\$1,000 genome a reality.

Producing genomic data is now relatively easy, but analysing these data is not. For precision medicine to fulfil its potential we need to identify genetic variation between individuals, and then work out which variants have a role in disease.

Human genomes are very similar, but the 0.1% difference between them still leaves millions of variations between individuals. Most such variations have little or no effect. Working out whether a particular deviation from the reference genome is important, and how, is complex and time consuming. It has become the crucial bottleneck in the precision-medicine process.

Drawing the connections between genetic variants and disease is largely the work of bioinformatics. Conventionally, the computer software used has been written and shared by academics. But as the production of genomic data has exploded, commercial firms have increasingly started to offer their own software. This growing market is evident to anyone who attends major genetics conferences. Three or four years ago, just a handful of these companies exhibited; now there are dozens.

The appeal of commercial bioinformatics packages is obvious. They are relatively simple to use, with well-designed interfaces that allow even non-experts to process complex genomic sequence information. Some commercial software streamlines the whole process, from sequencing to analysis and interpretation. The companies guarantee technical support, which is not always available with open-source software.

However, there is a major problem. Companies are generally unwilling to reveal how their software works: they do not wish to disclose the methods and data used to construct the algorithms or the details of performance. It is impossible to check the programs' quality and to compare them. Companies are selling a pig in a poke.

In such tools, the details really matter. Short segments of sequence (reads) must be joined to build complete genomes. This is difficult, and fast sequencing methods have quite high error rates, which must be taken into account when software flags possible variants. The more overlapping reads that a sequencing project includes, the better the results. Typically the coverage is in the tens, but in really deep sequencing it can be in the thousands.

Once possible variants are identified, a different set of techniques is used to filter and sort them, and then to annotate them to suggest

possible clinical relevance. There is no single correct way to do this, and various academic groups have produced distinct tools that all perform these tasks slightly differently. That is why the output alone — the variants and their link to disease — is not sufficient to judge their clinical relevance. We must know how the result was obtained and how the raw data were processed.

Academics are up-front about this, and are happy to show their working. This allows comparison, and a number of studies have been published in which the performance of several methods has been checked against independent benchmark data sets. These studies allow end-users to select the most suitable tool and get an idea of how reliable it is. This information should be included when data are published, especially if it has a direct clinical relevance. The journal *Human Mutation* demands it for studies that use and develop these tools.

At present, it is impossible to check the performance of commercial software in this way. I have asked companies to give me the relevant details, but they have refused. They all say that their method is the best, but offer no way for customers to verify that. As the market grows for these commercial packages (many of which, ironically, are based on open-source academic programs), so will the scale of the problem.

The way to sort this out is to test each of the different commercial programs with established benchmarks — data sets with known variant outcomes. But even if I were to buy a licence to use each of them (and these are not cheap), I would still be unable to do the comparison. The algorithms that drive such software are often developed

using the same data sets. To make the tests fair, we need to know how the algorithm was trained, so as to avoid using the same variants for both training and testing. This is something that the companies are unwilling to reveal.

Companies expect users to accept their 'black-box' solutions without knowing anything about the algorithm, training details, data sets used, method performance and use of benchmark data. This is not acceptable. Research must be based on openness and full accounts of the tools used.

Precision medicine must be evidence-based medicine. And evidence-based medicine is exactly what the name says. I understand that companies need to keep some trade secrets, but disclosing the information I discuss here will not jeopardize their competitive edge. These are details that we as the community have to demand if companies want to sell their products and services to us. ■

Mauno Vihinen is professor of medical structural biology at Lund University, Sweden.
e-mail: mauno.vihinen@med.lu.se

RESEARCH MUST BE
BASED ON
OPENNESS
AND
FULL
ACCOUNTS
OF THE TOOLS USED.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/elxesv

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

ASTRONOMY

Spots spotted on Vega star

One of the brightest stars in the night sky seems to have surface structures called starspots — a surprising finding for this particular star.

Torsten Böhm at the University of Toulouse in France and his colleagues used a telescope at France's Haute-Provence Observatory to look at Vega, a well-studied star that is roughly double the mass of the Sun. They found evidence of many faint spots: structures caused by magnetic-field changes that slightly alter the temperature in these areas. Vega is an 'A-type' star, a group thought to be incapable of generating magnetic fields and hence these spots.

The starspots could be linked to a weak surface magnetic field that was detected from Vega in 2009, the authors say.

Astron. Astrophys. 577, A64 (2015)

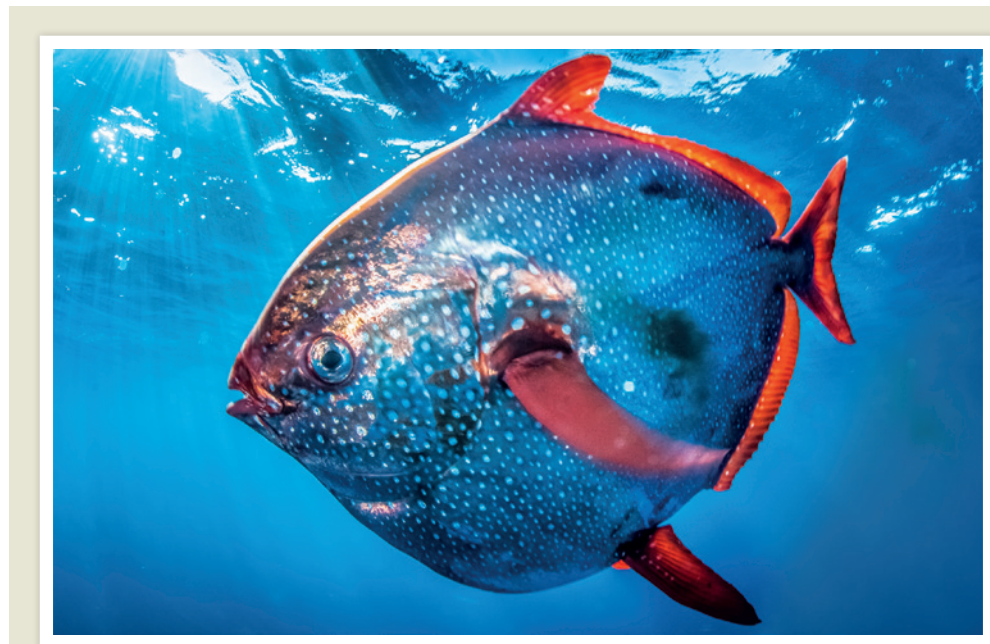
NEUROSCIENCE

A way to regrow nerve fibres

Injured neurons in fruit flies and mice regrow better when the activity of RtcA, an RNA-processing enzyme, is reduced.

Permanent damage to the central nervous system can occur when injured nerve cells fail to regenerate their axons — the long, impulse-transmitting part of the nerve cell. Yuh Nung Jan at the University of California in San Francisco and his colleagues screened fruit flies (*Drosophila melanogaster*) and found that severed axons regrew more effectively in mutant flies with reduced activity of RtcA.

When the enzyme was overexpressed, axons were regenerated less often and were much shorter than in normal



RALPH PACE

ANIMAL PHYSIOLOGY

Fish keeps warm in cold waters

A fish is able to maintain a warm body temperature in deep, cold waters.

Some species such as tuna can keep parts of their bodies warm, but Nicholas Wegner of the National Marine Fisheries Service in La Jolla, California, and his colleagues report that the deep-swimming opah (*Lampris guttatus*; pictured) can make its entire body — including the heart — warmer than its environment by

3–6 °C. They measured the temperature and studied the anatomy of 22 opahs captured off the coast of California at depths of 50–300 metres. They found that the animal generates heat by flapping its pectoral fins and retains it using specialized blood-vessel structures in the gills.

This warmth probably boosts the power output of the fish's muscles, the authors say.

Science 348, 786–789 (2015)

flies. Similar results were seen in rat cells and in mice.

Altering the activity of RtcA or other molecules that it regulates could offer treatments for nervous-system injuries, the authors suggest.

Nature Neurosci. <http://doi.org/4m3> (2015)

OPTICS

Iron atoms slow down X-rays

Researchers have made an X-ray beam travel 10,000 times slower than the speed of light — an effect seen before only

for visible light.

Physicists have previously slowed light waves to a crawl and even stopped them by controlling the transparency of the medium through which the light passed — usually an ultracold gas of atoms such as sodium. They did this by tuning the interaction of light with the electrons in the gas. Now, a team led by Jörg Evers of the Max Planck Institute for Nuclear Physics in Heidelberg, Germany, has seen a similar effect by letting X-rays from a synchrotron interact with the nuclei of iron atoms, rather than with their electrons.

Controlling X-rays in this way could be useful for high-resolution imaging and other applications.

Phys. Rev. Lett. <http://dx.doi.org/10.1103/physrevlett.114.203601> (2015)

CANCER

Organoids mimic tumours

Human cancer tissue that is grown into 'organoids' in the laboratory could be used to test drug responses and to personalize therapy.

Organoids are 3D cultures

of cancerous cells that better represent the composition of a tumour in the body than cancer-cell lines, according to Mathew Garnett at the Wellcome Trust Sanger Institute in Hinxton, UK, Hans Clevers at the Hubrecht Institute in Utrecht, the Netherlands, and their colleagues. They built a small bank of 22 tumour organoids using samples from 20 people with colon cancer, and tested the effects of 83 cancer drugs on the cultures. They found correlations between the activity of specific genes and responses to particular drugs.

Some organoids were also uniquely sensitive or insensitive to certain compounds, so the approach might one day be used to tailor treatments for individuals.

Cell 161, 933–945 (2015)

ASTRONOMY

Quasar quartet in galactic nursery

Astronomers have discovered a massive cluster of four quasars — a rare find of galaxies just being born.

Quasars are young, bright galaxies powered by supermassive black holes and are hard to find because this youthful period is brief. Using the W. M. Keck Observatory in Hawaii, Joseph Hennawi of the Max Planck Institute for Astronomy in Heidelberg, Germany, and his colleagues found the quasars (pictured, indicated by arrows) at the heart of one of the largest known nebulae — clouds of gas that, if large enough, can give birth to new galaxies. The quasars are illuminating the surrounding gas and

are probably evolving into a massive galaxy cluster.

This rare grouping, together with the size of the nebula, suggests that gas in protogalactic clusters might be cooler and denser than was thought.

Science 348, 779–783 (2015)

CHEMICAL BIOLOGY

Fish makes its own sunscreen

Zebrafish have the genes needed to synthesize a compound that can provide protection from ultraviolet radiation.

Such chemicals have been found in fish but it was thought that they came from their diet or from microbes that live in or on the animals. Taifo Mahmud at Oregon State University in Corvallis and his colleagues previously analysed fish genomes and discovered genes involved in making these compounds. They then studied the embryos of zebrafish (*Danio rerio*) and found that their extracts contained the sunscreen compound gadusol. The team inserted the zebrafish genes for gadusol production into yeast (*Saccharomyces cerevisiae*), which produced milligrams of the compound.

Yeast could be harnessed to make large quantities of the UV protectant, the authors say.

eLife 4, e05919 (2015)

ECOLOGY

Rare bees barely benefit ecosystem

The sheer number of the most common species in an ecosystem — rather than the level of biodiversity — determines how much the system benefits people.

Conservationists have argued that biodiversity supports ecosystem services such as crop pollination. To separate out the effects of species richness from species abundance, Rachael Winfree of Rutgers University in New Brunswick, New Jersey,

SOCIAL SELECTION

Popular topics on social media

Fruit-fly paper has 1,000 authors

Author lists have grown lengthy in many fields of science, but when a *Drosophila* genomics paper was published with more than 1,000 authors, it sparked discussion online about the meaning of authorship. The paper, published in the journal *G3: Genes Genomes Genetics*, names 1,014 authors, with more than 900 undergraduate students among them. Zen Faulkes, an invertebrate neuroethologist at the University of Texas–Pan American in Edinburg, questions on his blog whether every person made enough of a contribution to be credited as an author (see go.nature.com/8rffl7). But the paper's senior author, geneticist Sarah Elgin at Washington University in St. Louis, Missouri, says that large collaborations with correspondingly large author lists have become a fact of life in genomics research. "Putting together the efforts of many people allows you to do good projects," she says.

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/gssjdg

Genes Genomes Genet. 5, 719–740 (2015)



and her colleagues used an equation from evolutionary biology to analyse wild-bee pollination of fruit crops.

The team counted thousands of individual bees from as many as 56 different species in fields of watermelons, blueberries and cranberries, as well as the average number of pollen grains they deposited on flowers. They calculated that pollination was dominated by a few common bee species.

Loss of rare species would not change pollination rates much, but reductions in the number of common bees would make a huge difference, the authors say.

Ecol. Lett. <http://doi.org/4m5> (2015)

PALAEOLOGY

Gut microbes give good fossils

Gut microbes are the main driver of tissue decay when animals die, and were probably important for preserving

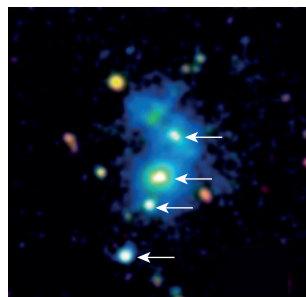
soft-tissue anatomy in fossil animals.

Philip Donoghue at the University of Bristol, UK, and his colleagues studied the brine shrimp (*Artemia salina*; pictured left) and monitored its decay (pictured, middle and right) under various conditions. They found that soon after death, the shrimp's gut wall breaks open and bacteria spill out into the body cavity. The bacteria form sticky aggregates, or biofilms, that gradually replace shrimp tissue and contain mineral deposits, as revealed by microscopy. This mineralization is a key step in tissue preservation in fossils.

Evolution of the gut led to an explosion in both animal diversity and the abundance of fossils, the authors say.

Proc. R. Soc. B 282, 20150476 (2015)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch



SEVEN DAYS

The news in brief

FUNDING

WHO crisis fund

The director-general of the World Health Organization, Margaret Chan, outlined plans for a US\$100-million fund on 18 May to help the agency to respond to global health emergencies. The fund, announced at the annual World Health Assembly in Geneva, Switzerland, will be filled by voluntary donations. The measure is one of many that WHO member states are discussing in light of what the agency has acknowledged were shortcomings in its response to the Ebola outbreak (Y.-A. de Montjoye *et al. Sci. Rep.* 3, 1376; 2013).

Australian grant cut

Australia's government will remove Aus\$263 million (US\$211 million) from university grants over the next three years to keep key national research facilities running, according to its 12 May budget. The cash is to save the National Collaborative Research Infrastructure Strategy (NCRIS), which employs 1,700 staff across 27 facilities, many of which faced closure earlier this year. Researchers persuaded the government to provide Aus\$300 million to save the NCRIS, but the budget now shows that this money will mostly be

NUMBER CRUNCH

5,154

The record-breaking number of authors on a single paper, published in *Physical Review Letters* on 14 May.

Phys. Rev. Lett. 114, 191803 (2015)



NORBERT WU/MINDEN PICTURES/CORBIS

Polar Code to curb pollution from vessels

Ships plying Arctic and Antarctic waters face specific environmental regulations for the first time, after the International Maritime Organization agreed rules to combat polar pollution on 15 May. The environmental provisions are designed to prevent pollution

from oil, sewage and rubbish from vessels, and will begin coming into force in 2017. The rules are an addition to the 'Polar Code', which was adopted in 2014 as the first set of standards specifically regulating polar shipping (see go.nature.com/xhsanz).

diverted from university grants. Overall, researchers have grudgingly accepted the science budget, which is flat relative to last year's, as a reprieve after years of funding cuts. See go.nature.com/h62vub for more.

Antibiotic fund

Drug companies should be offered lump-sum payments as a reward for developing new antibiotics in the face of growing drug resistance, according to a review commissioned by the UK government. The review, led by economist Jim O'Neill, says that companies would get an immediate return on their investment, rather than waiting years for an antibiotic to become widely used — which often happens near the end of

its patent. A global incentive system could cost as little as US\$16 billion over 10 years. O'Neill's team also proposes a global fund of \$2 billion for basic research on antibiotic development.

BUSINESS

Drugs super-spend

The number of people in the United States who spend more than US\$100,000 a year on medicines tripled to 139,000 last year, compared with 2013. The figure was published on 13 May in a report by Express Scripts, a prescription-management service headquartered in St Louis, Missouri. Hepatitis C drugs, cancer treatments and specially mixed formulas called 'compounded' medicines were

the main drivers of the rise. More than 500,000 people spent at least \$50,000 on drugs in 2014, compared with 352,000 in 2013 — a 63% rise. Prescriptions cost Americans \$374 billion in 2014. The rise in spending was the largest in a decade.

Science spin-offs

The University of Oxford, UK, is raising a £300-million (US\$470-million) fund to commercialize ideas from its science departments, it announced on 14 May. The money, towards which high-profile investors have already committed £210 million, will be channelled through Oxford Sciences Innovation, a company that will provide both capital and advice for the university's spin-off

JOSE NUÑEZ-MINO

companies. Oxford vice-chancellor Andrew Hamilton said that the venture would significantly expand the funding and commercial expertise available to the university's research teams.

RESEARCH

Peer-review log

Researchers can now use their ORCID profiles to keep a record of their peer-review activities. ORCID (Open Researcher and Contributor ID), a global non-profit organization that gives each researcher a unique 16-digit identifier and web page to log their scientific contributions, announced on 18 May a standardized format for recording peer reviews of manuscripts, grant applications and conference abstracts. Some publishers said that they would adopt the standard. See go.nature.com/8pvt4y for more.

James Bond rodent

A new rodent has been found on the Caribbean island of Hispaniola. The animal (**pictured**) is one of just eight remaining species of hutia — guinea-pig-like rodents native to the Caribbean — and is named James Bond's hutia (*Plagiodontia aedium bondi*), not after the fictional spy, but after the US naturalist



who inspired Ian Fleming's creation. Bond identified a barrier between ecosystems in southern Haiti, known as Bond's Line, which marks the boundary between the latest hutia and its closest relatives. The team that discovered the species, led by Samuel Turvey at the Zoological Society of London, warns that it is already endangered.

Summer bee losses

Losses of honeybee colonies in the United States have increased substantially over the past year. They are now more than double the level generally considered sustainable. The rate of colony loss in winter 2014–15 was 23.1% — below the rolling 9-year average of 28.7% — but summer losses exceeded winter ones for the first time, reaching 27.4%. The overall loss for April 2014 to April 2015 was 42.1%. Beekeepers consider a loss rate of 18.7% per year economically

acceptable, according to the US Department of Agriculture, which conducted the survey with non-profit partners.

El Niño to persist

El Niño conditions in the tropical Pacific Ocean are likely to persist throughout 2015, forecasting agencies have said. The gathering of warm waters in the east and central Pacific was first reported in March by the US National Oceanic and Atmospheric Administration (NOAA). In its forecast on 14 May, NOAA estimated that there is a 90% chance that the El Niño will continue into the Northern Hemisphere summer, and an 80% chance that it will persist until the end of 2015. El Niño conditions have been linked to extreme weather around the globe.

POLICY

EU science panel

The European Commission announced plans on 13 May to install a new scientific advice system to aid its policy decisions. The panel will replace the post of chief scientific adviser, which was created in 2012 but abolished in 2014 after it failed to match the complex needs of the commission. The commission will now recruit a high-level group of seven internationally renowned researchers, and

COMING UP

24–28 MAY

Planetary and Earth scientists meet in Chiba for this year's meeting of the Japan Geoscience Union.

go.nature.com/wg4j73

draw on scientific expertise in national academies and other learned bodies in European Union member states. The system will be operational by the autumn, the commission says. See go.nature.com/vlm19i for more.

Non-GM food stamp

US regulators plan to offer a 'non-GM' certification service to verify food manufacturers' claims that products are free of genetically modified (GM) ingredients. The programme was described by agriculture secretary Tom Vilsack in a letter to US Department of Agriculture (USDA) employees, but has not been publicly announced. The USDA already verifies other food-marketing claims for a fee, and Vilsack noted that companies have expressed interest in obtaining the non-GM verification.

EVENTS

Genome standard

The US National Institute of Standards and Technology (NIST) has created the first reference standard for validating DNA tests. On 14 May, NIST announced that it will sell tubes containing 10 micrograms of the reference genome — that of a woman with European ancestry — for US\$450 each. The material is intended for laboratories and companies to ensure the accuracy of their genetic-sequencing tests for diseases or personalized treatments.

➔ NATURE.COM

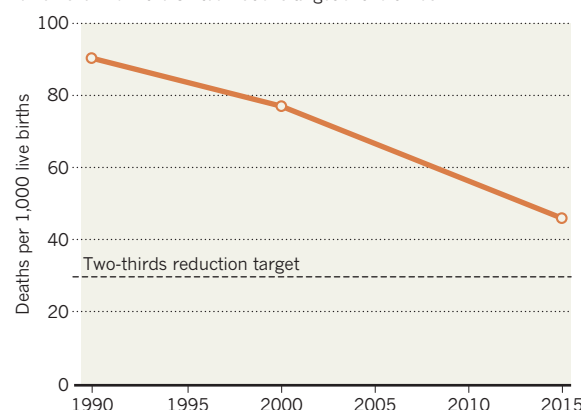
For daily news updates see:
www.nature.com/news

TREND WATCH

The world has made huge progress in some of the health-related goals set by the United Nations in 1990, said a report from the World Health Organization on 13 May. Although the 194 member states have together slashed child mortality by half, the world will not meet the Millennium Development Goal of cutting it by two-thirds by the end of this year. Much of the fall is from reducing deaths caused by pneumonia, diarrhoeal diseases, measles and malaria.

CHILD DEATHS FALL

Under-five mortality has halved since 1990, but the reduction remains far from the United Nations target of two-thirds.



SOURCE: WORLD HEALTH ORGANIZATION

NEWS IN FOCUS

POLICY US Congress debates crack down on patent trolls **p.270**

NUCLEAR PHYSICS Atomic dream machine edges closer to reality **p.272**

POLITICS Russian science foundation fears for its future **p.273**



REPRODUCIBILITY Why so many antibody results cannot be repeated **p.274**

PAOLO PELLEGRIN/MAGNUM



Poppies may soon be joined by bioengineered yeast as a source of opiates.

SYNTHETIC BIOLOGY

Engineered yeast paves way for home-brew heroin

Advance holds potential for better opiate painkillers — but raises concerns about illicit use.

BY RACHEL EHRENBERG

Biotecnology is about to make morphine production as simple as brewing beer. A paper published on 18 May in *Nature Chemical Biology*¹ reports the creation of a yeast strain containing the first half of a biochemical pathway that turns simple sugars into morphine — mimicking the process by which poppies make opiates. Combined with other advances, researchers predict that it will be only a few years — or even months — before a single engineered yeast strain can complete the entire process.

Besides giving biologists the power to tinker with the morphine-production process, the advance could lead to more-effective, less addictive and cheaper painkillers that could be brewed under tight controls in fermentation vats. At the same time, it could enable widespread, localized production of illegal opiates such as heroin, increasing people's access to such drugs. Recognizing that danger, the synthetic biologists behind the work have already opened a discussion of how to prevent the technology's misuse without hampering further research.

"It's easy to point to heroin; that's a concrete

problem," says bioengineer John Dueber of the University of California, Berkeley, who led the latest research. "The benefits are less visible. They are going to greatly outweigh the negative, but it's hard to describe them."

Over the past decade, several research teams have tried to coax microbes, including the baker's yeast *Saccharomyces cerevisiae* and the lab-workhorse bacterium *Escherichia coli*, into making plant-derived drugs. The anti-malarial drug artemisinin, originally derived from sweet wormwood (*Artemisia annua*), is now produced commercially in yeast.

The opium poppy (*Papaver somniferum*), ▶

► as the only commercial source of morphine and opioid painkillers such as oxycodone and hydrocodone, is an obvious target for bioengineering. The crop must be grown in highly regulated conditions, in only a few countries. Outside those boundaries, in places such as Afghanistan, it is grown to supply the illegal heroin trade. Producing opiates in industrial facilities from yeast could eliminate the need for the tightly controlled legal plant-production chain.

But the opiate-synthesis pathway is long — roughly 18 steps — and biochemically complex. Because there is no whole sequenced genome for the opium poppy, identifying the enzymes that catalyse the synthesis reactions has been difficult. So bioengineers have looked for enzymes in other plants, and even in humans and insects, that could carry out the desired reactions when inserted into a microbe's genome. They have also developed techniques to make those enzymes more efficient by mutating the genes that encode them and selecting for mutations that increase the output of desired products. But so far, no one has been able to engineer the whole process into a single organism.

ASSEMBLY REQUIRED

Dueber and colleagues' work does not reach that goal. But it demonstrates that, given the right genes and biochemical machinery, yeast can convert glucose into the intermediate

compound (S)-reticuline — the first half of the poppy's morphine-production pathway.

Together with a similar paper² published in *PLoS ONE* on 23 April that covers the second half of the pathway, and a PhD thesis (see go.nature.com/kwgc8n) that identifies the cellular machinery to join the two halves, all of the pieces are now in place to make opiates in yeast.

"I don't want to undersell how much work there still is to do, but I don't want to undersell how short that work is," Dueber says. Even when the entire apparatus has been incorporated into a single strain of yeast, efforts will still be needed to make the fermentation processes efficient. In theory, once that work is done, anyone who could obtain the engineered yeast strain would be able to make morphine in a process that is no more complicated than home-brewing beer.

For that reason, Dueber and his colleagues shared their research before publication with biotechnology-policy specialists Kenneth Oye, of the Massachusetts Institute of Technology (MIT) in Cambridge, and Tania Bubela, of the University of Alberta in Edmonton, Canada. With MIT political scientist Chappell Lawson, Oye and Bubela have written a Comment article³ in *Nature* (see page 281) that calls for proactive examination of the risks and benefits of engineering organisms to make compounds

that are both useful and dangerous. They urge drug and biosecurity regulators, law-enforcement agencies, scientists and public-health officials to come together to develop safeguards that minimize risk without quashing research.

"From the perspective of law enforcement, this is a new technology that could be abused with negative consequences," says Lawson, who spent 18 months as an adviser to the commissioner of US Customs and Border Protection. "I don't think anyone wants millions more opiate addicts."

BEET ROUTE

Clever biochemical prospecting helped Dueber and his colleagues to advance their research faster than expected. The scientists needed an enzyme that would catalyse the initial reaction in the morphine assembly line, a step that has been a stumbling block in making the process work in yeast. That stage turns tyrosine, an amino acid that yeast naturally produces in abundance, into the molecule L-DOPA. But most enzymes that catalyse that reaction go on to turn L-DOPA into an unwanted by-product. So William DeLoache, a bioengineer in Dueber's lab, took such an enzyme from sugar beet (*Beta vulgaris*) and systematically mutated it until it would carry out only the reaction from tyrosine to L-DOPA (see 'Pathway potholes'). Adding more enzymes, including one from a soil bacterium and one that the team discovered in the opium poppy, completed the first half of the opiate assembly line.

The implications for opiate production are one thing, but researchers say that they are most excited by the prospect of taking bits and pieces of plant pathways to create entirely new molecules. "A plant goes from A to Z without stopping at something valuable in between," says biologist Vincent Martin of Concordia University in Montreal, Canada, who contributed to the development of both the first and second halves of the yeast production pathway. "We're not restricted to what evolution has restricted to a single plant — we can mix and match."

That is the real promise of synthetic biology, says Dueber: not just cutting the cost and time it takes to make known plant compounds, but tinkering with the processes that a plant uses to make what he calls "unnatural natural products". These have the potential to be extremely beneficial, but that value risks being overshadowed by the spectre of illicit yeast-based heroin production.

"This work has very interesting and important implications, but there are regulatory gaps," says Oye. "The question is, can regulators be nimble and figure this out before people finish the work?" ■ [SEE COMMENT P.281](#)

1. DeLoache, W. C. *et al.* *Nature Chem. Biol.* <http://dx.doi.org/10.1038/nchembio.1816> (2015).
2. Fossati, E., Narcross, L., Ekins, A., Falgout, J. P. & Martin, V. J. J. *PLoS ONE* **10**, e0124459 (2015).
3. Oye, K., Bubela, T. & Lawson, J. C. H. *Nature* **521**, 281–283 (2015).

PATHWAY POTHOLES

The proof is in the pigment

The effort to manufacture opiates in yeast has long been stymied by one step early in the process by which poppies make such drugs. This reaction converts the amino acid tyrosine into L-DOPA, which can then be converted into dopamine, a compound needed in large quantities for opiate production. The enzyme that opium poppies use for this step has not yet been identified. Although enzymes from other plants carry out the reaction, they also immediately convert L-DOPA into dopaquinone, an unwanted by-product that many organisms use to make the pigment melanin.

To solve this problem in the yeast pathway, bioengineer William DeLoache and colleagues at the University of California, Berkeley, inserted¹ a sugar-beet enzyme into yeast that converts tyrosine into L-DOPA. Then they induced mutations in the yeast cells to produce multiple variants of the enzyme, hoping that one would churn out L-DOPA without also catalysing the reaction that makes dopaquinone. To spot the hoped-for enzyme, they inserted into the mutated yeast another plant enzyme, from

the four o'clock flower (*Mirabilis jalapa*). That enzyme produces the bright-orange pigment betaxanthin when L-DOPA is present, and a purple one in response to the unwanted dopaquinone. This revealed, simply by colour, which mutants most efficiently made L-DOPA from tyrosine but did not make dopaquinone.

After screening nearly half a million mutants, the team hit on an enzyme that produced one-fifth the amount of violet pigment compared to the yeast containing the original beet enzyme. This variant also made 3.7 times more orange betaxanthin. The researchers then removed the pigment-making enzymes from the mutant, replacing them with an enzyme from the soil microbe *Pseudomonas putida*, which efficiently converts L-DOPA into dopamine. They also found a previously unknown enzyme from opium poppies that converts dopamine into the next morphine precursor, norcoclaurine. By adding four more known enzymes to catalyse subsequent reactions, the team assembled the entire first half of the opiate-production pathway in yeast. **R.E.**



A Peruvian woman with poultry received as part of an aid programme to encourage self-sufficiency.

DEVELOPING WORLD

Short-term aid has long-term impact

Experiment across six nations shows that two-year interventions help to lift people out of extreme poverty.

BY DECLAN BUTLER

Giving some of the world's poorest people a two-year aid package — including cash, food, health-care services, training and advice — improves their livelihoods for at least a year after the support is cut off, according to the results of an experiment involving more than 10,000 households in 6 countries.

The poverty intervention had previously been trialled successfully in Bangladesh, and the study's researchers say the latest work shows that the approach works in other cultures, too. "We finally have truly credible evidence that a programme for the poorest of the poor can really help them meaningfully reduce their poverty," says Dean Karlan, an economist at Yale University in New Haven, Connecticut, and a co-author of the study, reported last week in *Science*¹. "Until now, we haven't really been able to go to a government outside Bangladesh and say, we're confident this works."

Ethiopia, one of the countries in the latest trial, is planning to continue the intervention and scale it up to cover around 3 million people, says Karlan. Pakistan and India are also considering scaling up interventions.

Outside experts are more cautious, but

still impressed, particularly because the work was done as a randomized controlled trial in which people were randomly assigned to either an intervention or a control group, much in the way that drugs and vaccines are tested. Most poverty interventions have failed to show sustainable benefits in such trials, so the effectiveness of the programme justifies countries considering the strategy, says Jonathan Morduch of New York University, who studies micro-finance and poverty.

"A programme for the poorest of the poor can really help them meaningfully reduce their poverty."

Randomized controlled trials to test poverty interventions were developed over the past decade by the Abdul Latif Jameel Poverty Action Lab (J-PAL), at the Massachusetts Institute of Technology in Cambridge (see *Nature* **493**, 462–463; 2013), and Innovations for Poverty Action, a New Haven-based non-profit organization founded by Karlan that coordinated the latest study.

The programme tested in the study uses the 'graduation model', which aims to graduate people out of extreme poverty. It was invented in Dhaka by the Bangladesh Rural Advancement

Committee (BRAC), one of the world's largest non-governmental development organizations. More than 1 billion people in the world live on less than US\$1.25 per day, but the graduation model targets the hundreds of millions who live on less than 70 cents per day. These people are mostly rural women and slum dwellers who are often dependent on aid to survive.

By 2011, BRAC had reached some 400,000 households in Bangladesh, and a 2013 report² of a randomized trial concluded that its programme was highly effective. The latest study tested whether the intervention would work elsewhere. Households were given assets such as goats, sheep or chickens to start farming, or the means to open a shop, and were supported with food, cash, a savings account and access to health care while they were getting their activity up and running. Coaches visited regularly over two years to offer advice — such as how to manage money — and keep people on track.

Overall, one year after the intervention stopped, the experiment had produced a 14% increase in assets and a 96% increase in savings, compared with those in similar groups of people not enrolled in the programme, the paper says. "Effects often fade over time, so seeing results persist for a year is already quite impressive," says Morduch. It shows that a coordinated short-term intervention can put very poor people on the first rung of the ladder to escape from extreme poverty.

NO PANACEA

Although the intervention was successful in Ethiopia, Ghana, India, Pakistan and Peru, it failed in Honduras. There, poor households were mostly given imported chickens, many of which caught local diseases and died. Morduch also worked on another graduation-model study in rural southern India, published in March³, that failed to show any benefits. Residents there turned out not to be keen to become farmers, Morduch says, and most ended up selling the livestock that they were given to take up paid labour.

The intervention is also not cheap. Costs per household ranged from \$1,455 in India to \$5,962 in Pakistan, although they were offset by positive returns on investment ranging from 133% in Ghana to 433% in India. The researchers hope to cut costs in future by scaling back the experiment's more expensive components, such as training.

The graduation model is no cure-all: historically, the biggest reductions in extreme poverty have resulted from larger economic growth, notes Karlan. But trickle-down economic improvements will not end widespread extreme poverty any time soon, he says, so there is a pressing need for bottom-up interventions. ■

1. Banerjee, A. *et al.* *Science* <http://dx.doi.org/10.1126/science.1260799> (2015).
2. Bandiera, O. *et al.* *IZA Discussion Paper No. 7386* (2013); available at <http://go.nature.com/9pedna>
3. Bauchet, J., Morduch, J. & Ravi, S. *J. Dev. Econ.* **116**, 1–16 (2015).



Transgenic mice were the subject of a recent patent-licensing fight in the United States.

POLICY

Congress seeks to quash patent trolls

Revised legislation would spare universities from being penalized in the same way as unscrupulous companies.

BY HEIDI LEDFORD

Predatory 'patent trolls' could soon find it harder to operate in the United States. Legislation to curb frivolous patent lawsuits has regained momentum after lawmakers in the US Senate added a provision to stop university patent holders from being penalized along with the trolls.

The process is moving quickly. The Senate Judiciary Committee plans to vote on the bill by the end of the month, readying it for a final Senate vote this summer, and the House

of Representatives' Judiciary Committee is likely to vote this week on a similar measure. That gives observers optimism that Congress will finally enact patent-troll legislation after a failed effort last year. "The Senate version really does seem to be hitting some sort of sweet spot," says Arti Rai, co-director of the Duke Law Center for Innovation Policy in Durham, North Carolina.

Patent trolls are 'non-practising entities' that accumulate patents with no intention of turning the inventions into marketable products. Many of these firms exist solely to enforce

their patents, threatening other companies with lawsuits if they are not paid handsome licensing fees. The legal strategy is often a low-risk endeavour, because many patent trolls are shell corporations that are only loosely affiliated with larger firms — and so do not have the financial assets that would support large awards to opponents should they lose a suit.

Congress, urged on by lawsuit-weary high-technology companies and the administration of US President Barack Obama, is trying to fight these kinds of trolls. Non-practising entities filed 63% of all US patent-infringement lawsuits in 2014, and cost operating companies an estimated US\$12.2 billion in legal fees, settlements and judgments, according to RPX, a consultancy in San Francisco, California.

To fight the shell corporations, in February, several members of the House introduced a bill that would hold all owners of a patent liable for opponents' legal fees if the owners lose a patent-enforcement suit. That would mean that the parent firm could be compelled to pay, making nuisance lawsuits more costly — and a higher risk — for trolls.

But the requirement would also have extended to universities, and could have threatened their ability to defend their own patents. Universities, too, are non-practising entities because they patent inventions but often do not directly commercialize them. Instead, they charge other companies for the right to turn those patents into products. The Senate is addressing that problem by introducing a bill that exempts institutions of higher education.

But some legal scholars have raised eyebrows at the carve-out. "Universities aren't exactly coming to this argument with clean hands," says Tania Bubela, who analyses health and biotechnology law at the University of Alberta in Edmonton, Canada. Universities sometimes license their patents to other non-practising entities, including some that are widely considered to be trolls (see *Nature* 501, 471–472; 2013). But, she says, the Senate legislation only skims the surface of the problem. "They're not addressing the root issue, and that is the mess that is patent examination at the US Patent and Trademark Office," she says.

The US patent office is often criticized for granting patents too readily, resulting in a gnarly — and growing — thicket of patents (see 'Patent pile-up'). The result is that

ROBERT F. BUKATY/AP



**MORE
ONLINE**

TOP STORY



Early Europeans may have had Neanderthal great-great grandparent
go.nature.com/6lawlj

MORE NEWS

- UK researchers fret about downgrading of science-minister role
go.nature.com/fyaj6k
- Indian ocean may be key to global warming 'hiatus' go.nature.com/k8sInf
- Model finds risky regions for human-animal diseases go.nature.com/jbhwm5

NATURE PODCAST



The oldest stone tools yet found, and the perks of sex ... for beetles.
nature.com/nature/podcast

SOURCE: USPTO

companies often struggle to discern when they are infringing intellectual property (see *Nature* **458**, 952–953; 2009). A series of court decisions has begun to address the problem, says Nicholson Price, a legal scholar at the University of New Hampshire in Concord. Foremost among them is a Supreme Court decision last year to limit patents on software, which has yielded a steady stream of district-court decisions to invalidate questionable patents (see *Nature* **507**, 410–411; 2014). The patent office has also created a process by which outside parties can challenge recently granted patents without resorting to litigation, which has helped to tighten patent standards (see *Nature* **472**, 149; 2011).

DROP IN THE OCEAN

Even so, says Robin Feldman, director of Institute for Innovation Law in San Francisco, California, there is a need for Congress to enact further patent reforms. “All of these measures

PATENT PILE-UP

The number of patents granted in the United States has risen sharply over the past few decades.



are needed,” she says. “There is no silver bullet.”

Universities might even benefit from the added protections. Although patent lawsuits against academic researchers are rare, they are legal. In 2010, a non-practising entity called the Alzheimer’s Institute of America (AIA)

in Sarasota, Florida, sued several institutions for infringing its patents on some transgenic mice used to study Alzheimer’s disease. One of the defendants was the University of Pennsylvania in Philadelphia, which had spun off a company to commercialize discoveries made using the mice. The AIA also sued the Jackson Laboratory, a widely used non-profit repository of research mice in Bar Harbor, Maine, and pressured the laboratory to relinquish a list of all researchers who had ordered the mice in question.

The case was dismissed in 2011 without the list being released, but the lawsuit’s legacy still lingers. Some researchers hesitate to share their transgenic mice for fear of putting themselves at risk, says Michael Sasner, who was in charge of the Jackson Laboratory’s Alzheimer’s resources at the time of the lawsuit. “The effect is that these mice aren’t being used to help develop drugs,” he says. “There’s got to be a better way.” ■

POLICY

UK slack on misconduct reports

Few universities follow guidelines to publish their records of investigations.

BY ELIZABETH GIBNEY

Just a fraction of universities in the United Kingdom have made public the extent of their investigations into research misconduct — even though all have been told that they should do so.

Since 2013, the United Kingdom’s major research funders have said that to receive grants, universities must adhere to a set of guidelines that recommend publishing annual summaries of their formal investigations into research misconduct.

But a survey on behalf of the UK Research Integrity Office (UKRIO, a national advisory body with no regulatory powers) has found that universities are falling short on this recommendation. It was presented at the UKRIO’s annual conference in London on 13 May.

The integrity guidelines are laid out in a document called *The Concordat to Support Research Integrity*, which was created in 2012 to counteract claims that the United Kingdom, which has no regulatory body covering research integrity, had an inadequate system of oversight to deal with research misconduct.

The survey contacted 44 universities that contribute funding to the UKRIO, and found that of 27 who responded, only one-third had published summaries of their investigations into research misconduct for 2013–14. Among another 44 randomly chosen institutions who

do not subscribe to the UKRIO, the figure was 7% — just 3 institutions. The UKRIO plans to publish the survey at a later date.

The 12 reports that had been published outlined a total of 21 investigations, of which 4 upheld the allegations of misconduct and 3 were ongoing. Of the 11 cases in which a type of misconduct was specified, 5 cases were investigations into plagiarism, 2 into falsification, 2 into questions of authorship, 1 into fabrication and 1 into breach of confidentiality.

It became clear at the conference that not every university had the same understanding of the concordat’s wording that institutions ‘should’ make their reports public. Not all took it to mean that reporting was mandatory; those that did included the University of Cambridge. “We didn’t know we were an outlier,” said Peter Hedges, head of the university’s research-operations office and a member of the UKRIO advisory board.

THE MEANING OF ‘SHOULD’

Survey author Elizabeth Wager, a freelance consultant and a member of the UKRIO’s advisory board, said that she too interprets the recommendation as a requirement. But

she understands universities’ reluctance to publish their data, she added. “Properly conducted misconduct investigations should be seen as a badge of honour, not something you’re embarrassed about. If there’s an increase in them, that might be a good thing. However that’s not always how the public perceives it, or the way it’s written up, so I can understand the caution,” she told the conference.

Institutions may also worry that their definitions of misconduct and formal investigation differ from those of other institutions, she said, so more guidance to ensure that universities are counting the same things would be helpful.

Even for the 12 reports that were published, finding information was not always easy, said Wager. In one case, she needed a login to access the published report; in another, the number of investigations was not stated.

Four universities’ reports stated that they had no formal investigations — which probably stemmed from differing definitions of what counts as an investigation, Wager said. “I think it’s completely improbable for big, research-intensive universities to say we have had no cases. It’s just not credible.”

Wager added that many universities she spoke to as part of the research said that they were in the process of putting together these reports, or were planning to do so next year. ■ [SEE EDITORIAL P.259](#)

“Properly conducted misconduct investigations should be seen as a badge of honour.”



Brookhaven National Laboratory in New York is a potential host for the Electron-Ion Collider.

NUCLEAR PHYSICS

Billion-dollar collider gets thumbs up

Proposed US electron-ion smasher wins endorsement from influential nuclear-science panel.

BY EDWIN CARTLIDGE

A machine that would allow scientists to peer deeper than ever before into the atomic nucleus is a big step closer to being built. A high-level panel of nuclear physicists is expected to endorse the proposed Electron-Ion Collider (EIC) in a report scheduled for publication by October. It is unclear how long construction would take.

The panel is the Nuclear Science Advisory Committee, or NSAC, which produces regular ten-year plans for the US Department of Energy (DOE) and the National Science Foundation. Its latest plan is still being finalized, but NSAC's long-range planning group "strongly recommended" construction of the EIC at a meeting last month, says NSAC member Abhay Deshpande, a nuclear physicist at Stony Brook University in New York. The EIC will almost certainly be formally endorsed in the NSAC report, he says. It must then be approved by the DOE, but most projects backed by the expert panel have come to fruition, he says.

The collider would allow unprecedented insights into how protons and neutrons are

built up from quarks and the particles that act between them, known as gluons.

The current leading facilities for studying quark-gluon matter are the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory in Upton, New York, and the Large Hadron Collider at CERN, Europe's particle-physics laboratory near Geneva, Switzerland. These facilities smash protons and heavy ions together to recreate the energetic conditions of the early Universe, when quarks and gluons existed as a plasma rather than in atomic nuclei. The EIC would collide point-like electrons with either protons or heavy ions, generating collisions that have a similarly high energy but are more precise and so can be used to study subatomic particles in detail.

In particular, the EIC would be ideal for studying an exotic state of matter that is made up entirely of gluons. The machine should also solve a puzzle about the proton that has baffled physicists for nearly 30 years. The proton has

a quantum-mechanical property called spin, but, strangely, the spins of its three constituent quarks add up to only about one-third of its own spin. The EIC would determine what makes up the difference: options include the spin of the proton's gluons, the angular momentum of its quarks or of the gluons from their orbital motion, or a mixture of all three.

"Until we have the EIC, there are huge areas of nuclear physics that we are not going to make progress in," says Donald Geesaman, a nuclear physicist at Argonne National Laboratory in Illinois, and the chair of NSAC.

The machine would not be built from scratch. One option is to add an electron-beam facility to RHIC — a plan that is estimated to cost about US\$1 billion and would depend on some as-yet-unproven technologies. Another is to add an ion accelerator and new collider rings to the Continuous Electron Beam Accelerator Facility at the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, which would cost about \$1.5 billion.

Deshpande hopes that the DOE will give the collider the thumbs up within a year of the NSAC plan's publication. Two or three more years would be needed to finalize the competing bids and choose one, meaning that construction could start in about 2020 and be completed five years later, he says.

Others say that this outlook is too rosy. The 2008 financial crisis led to a drop in science funding that forced NSAC to review its 2007 ten-year plan. A specially formed subcommittee concluded in 2013 that RHIC would have to shut down if funding for the DOE's Office of Nuclear Physics remained flat over the following five years. In fact, those funds have grown slightly, keeping RHIC in business, but the scare led to a more cautious approach this time around, says Geesaman. He points out that when the DOE and the National Science Foundation commissioned the ten-year plan, they specified that NSAC should consider what US physicists could achieve if funding remained flat, as well as how much support they would need to maintain a "world-leadership position".

Robert McKeown, deputy director for science at the Jefferson lab, thinks that limited funds might delay the start up of the EIC until at least 2030. And Michael Lubell, director of public affairs at the American Physical Society, questions whether it is feasible for the EIC to be built by the United States alone. He notes that the \$1.5-billion Long-Baseline Neutrino Experiment became an international project after a slimmed-down \$600-million version failed to pass scientific muster. "It is hard to see how to do this unless you get international buy-in," he says.

Deshpande thinks that the United States can go it alone. But he notes that collaborations at CERN and in China are also developing plans for electron-ion colliders and that the three groups are already exchanging ideas. ■

BROOKHAVEN NATIONAL LABORATORY/CC BY-NC-ND 2.0

POLITICS

Russia turns screw on science foundation

Ministry of Justice threatens to label Dynasty Foundation a 'foreign agent'.

BY QUIRIN SCHIERMEIER

Dmitry Zimin is not your typical Russian oligarch. Whereas some choose to pump their fortunes into yachts and football clubs, the 82-year-old former president of VimpelCom, one of Russia's largest telecommunication companies, turned to philanthropy — creating modern Russia's first private science-funding organization.

The Dynasty Foundation has earned Zimin prestigious government awards, but has now fallen out of favour with the Kremlin, putting its future at risk.

After a lengthy audit of the 13-year-old foundation's finances and funding activities, the Russian Ministry of Justice indicated in a preliminary report last month that Dynasty qualifies as a 'foreign agent', although it offered no justification for why. The designation relates to a controversial law passed by President Vladimir Putin in 2012 that requires non-governmental organizations that receive foreign funding and are deemed to be involved in vaguely defined "political activities" to register with the ministry as foreign agents.

The term is a loaded one to Russian ears, carrying Soviet-era connotations of espionage and treachery. Dynasty's executive director, Anna Piotrovskaya, says that its possible association with the foundation has deeply upset Zimin. Dynasty officials and Russian scientists fear that finalizing the label would seriously damage the reputation of the foundation and hamper its activities, and might even lead Zimin to shut it down.

So far, the ministry has labelled more than 50 human-rights groups, news agencies and environmental watchdogs as foreign agents, and issued warnings or brought charges against dozens of other groups for failing to register. Dynasty would be the first natural-sciences organization to be targeted, following at least two social-science research centres.

"The foreign-agent law is an outrageous attack on free speech," says Tanya Lokshina, who is the Russia programme director at Human Rights Watch in Moscow. "Sadly, hostility to anything foreign is as high in Russia as it has been since the height of the cold war."

The Dynasty Foundation is a force in Russian science. This year, it plans to spend some 435 million roubles (US\$8.9 million) on



Dmitry Zimin is deeply upset by the prospect of his foundation being labelled a foreign agent.

fellowships, summer schools and educational projects. Fellowship grants are available for molecular biologists, theoretical physicists and mathematicians at various stages in their careers: in 2014, Dynasty helped 288 physicists, 14 mathematicians, 21 biologists and 509 science teachers. Support from the foundation — up to 600,000 roubles (US\$12,000) per year for a postdoctoral researcher — has allowed hundreds of young Russian scientists to supplement their poor salaries, buy lab equipment and travel to meetings abroad.

FAIR PEER REVIEW

Many Russian scientists are shocked and outraged by the suggestion that Dynasty be designated a foreign agent. They hold the foundation in high regard for its cosmopolitan air — it requires all grant proposals to be submitted in English, for example — and its strictly merit-based funding criteria. These qualities help Russian science to flourish, something that the government has professed to be committed to, says Konstantin Severinov, a molecular biologist at the Skolkovo Institute of Science and Technology near Moscow, and a professor at Rutgers University in Piscataway, New Jersey. "For scientists who grew up in Russia, English-language proposals and fair peer review are very new experiences," he says. "These kinds

of things are really indispensable for bringing our scientists closer to their Western counterparts." They are a far cry from counting as the activities of a spy agency, he says.

Dynasty also seeks to foster free speech and scientific enlightenment. It supports the translation and publication of popular science books, such as physicist Stephen Hawking's *The Universe in a Nutshell* (Bantam, 2001). But although Dynasty has been careful to remain apolitical, says Lokshina, it might have been scuppered by its support of a Moscow museum and exhibition hall named after the famous Soviet dissident Andrei Sakharov.

Some, however, place the blame on an increasingly nationalistic and inward-facing government. Tightening the screws on science and education is consistent with the ongoing crackdown on liberal groups and ideas, says Fyodor Kondrashov, a group leader at the Centre for Genomic Regulation in Barcelona, Spain, who organizes Dynasty-funded summer schools in molecular and theoretical biology. "Russian society is undergoing hard-line changes," he says. "Where will it end?"

The popular summer schools are held every August in the academic city of Pushchino south of Moscow, and bring together talented high-school students, postdoctoral researchers and senior scientists from Russia and abroad. Experiments done there have produced several papers published in international journals, with students as co-authors. If the foundation is labelled a foreign agent, parents who back Putin, as most Russians do, might not want to send their children to the schools, and scientists who do the lecturing might also grow cautious. "Supporters of Putin and anyone fearful of trouble will stay far away," says Kondrashov.

As *Nature* went to press, the justice ministry had not announced its decision on the issue, which it was expected to do on 13 May. Members of Dynasty's board have scheduled a meeting on 8 June to discuss the ramifications of any ruling. The case is an example of problems that have been plaguing Russian science for years, says Valery Yakubovich, a sociologist at ESSEC Business School in Cergy Pontoise, France, including questionable accusations of espionage and political subversion against individual researchers. "Whatever happens," he says, "this is all very troubling." ■

BLAME IT ON THE ANTIBODIES

Antibodies are the workhorses of biological experiments, but they are littering the field with false findings. A few evangelists are pushing for change.

BY MONYA BAKER



In 2006, things were looking pretty good for David Rimm, a pathologist at Yale University in New Haven, Connecticut. He had developed a test to guide effective treatment of the skin cancer melanoma, and it promised to save lives. It relied on antibodies — large, Y-shaped proteins that bind to specified biomolecules and can be used to flag their presence in a sample. Rimm had found a combination of antibodies that, when used to ‘stain’ tumour biopsies, produced a pattern that indicated whether the patient would need to take certain harsh drugs to prevent a relapse after surgery. He had secured more than US\$2 million in funding to move the test towards the clinic.

But in 2009, everything started to fall apart. When Rimm ordered a fresh set of antibodies,

his team could not reproduce the original results. The antibodies were sold by the same companies as the original batches, and were supposed to be identical — but they did not yield the same staining patterns, even on the same tumours. Rimm was forced to give up his work on the melanoma antibody set. “We learned our lesson: we shouldn’t have been dependent on them,” he says. “That was a very sad lab meeting.”

Antibodies are among the most commonly used tools in the biological sciences — put to work in many experiments to identify and isolate other molecules. But it is now clear that they are among the most common causes of problems, too. The batch-to-batch variability that Rimm experienced can

produce dramatically differing results. Even more problematic is that antibodies often recognize extra proteins in addition to the ones they are sold to detect. This can cause projects to be abandoned, and waste time, money and samples.

Many think that antibodies are a major driver of what has been deemed a ‘reproducibility crisis’, a growing realization that the results of many biomedical experiments cannot be reproduced and that the conclusions based on them may be unfounded. Poorly characterized antibodies probably contribute more to the problem than any other laboratory tool, says Glenn Begley, chief scientific officer at TetraLogic Pharmaceuticals in Malvern, Pennsylvania, and author of a controversial

ILLUSTRATION BY NIK SPENCER/NATURE

analysis¹ showing that results in 47 of 53 landmark cancer research papers could not be reproduced.

A few scientists who have been burned by bad experiences with antibodies have begun to speak up. Rimm's disappointment set him on a crusade to educate others by writing reviews, hosting web seminars and raising the problem in countless conference talks. He and others are calling for the creation of standards by which antibodies should be made, used and described. And some half a dozen grass-roots efforts have sprung up to provide better ways of assessing antibody quality.

But it is too soon to call the cause a movement. "There are all these resources out there, but nobody uses them and many people aren't even aware of them," says Len Freedman, who heads the Global Biological Standards Institute, a non-profit group in Washington DC committed to improving biomedical research. "Most vendors have no incentive to change what's going on right now, even though a lot of the antibody reagents suck."

BUYER BEWARE

Take the example of Ioannis Prassas, a proteomics researcher at Mount Sinai Hospital in Toronto, Canada. He and his colleagues had been chasing a protein called CUZD1, which they thought could be used to test whether someone has pancreatic cancer. They bought a protein-detection kit and wasted two years, \$500,000 and thousands of patient samples before they realized that the antibody in the kit was recognizing a different cancer protein, CA125, and did not bind to CUZD1 at all². In retrospect, Prassas says, a rush to get going on a promising hypothesis meant that he and his group had failed to do all the right tests. "If someone says, 'Here is an assay you can use,' you are so eager to test it you can forget that what has been promised is not the case."

Most scientists who purchase antibodies believe the label printed on the vial, says Rimm. "As a pathologist, I wasn't trained that you had to validate antibodies; I was just trained that you ordered them."

Antibodies are produced by the immune systems of most vertebrates to target an invader such as a bacterium. Since the 1970s, scientists have exploited antibodies for research. If a researcher injects a protein of interest into a rabbit, white blood cells known as B cells will start producing antibodies against the protein, which can be collected from the animal's blood. For a more consistent product, the B cells can be retrieved, fused with an 'immortalized' cell and cultured to provide a theoretically unlimited supply.

Three decades ago, scientists who needed antibodies for their experiments had to make them themselves. But by the late 1990s, reagent companies had started to take over the chore.

Today, more than 300 companies sell over

2 million antibodies for research. As of 2011, the market was worth \$1.6 billion, according to global consultancy Frost & Sullivan.

DEVASTATING EFFECTS

There are signs that problems with antibodies are having broad and potentially devastating effects on the research record. In 2009, one journal devoted an entire issue to assessing the antibodies that are used to study G-protein-coupled receptors (GPCRs) — cell-signalling proteins that are targeted by drugs to treat various disorders, from incontinence to schizophrenia. In an analysis³ of 49 commercially available antibodies that targeted 19 signalling receptors, most bound to more than one protein, meaning that they could not be trusted to distinguish between the receptors.

The field of epigenetics relies heavily on antibodies to identify how proteins that regulate gene expression have been modified. In 2011, an evaluation⁴ of 246 antibodies used in epigenetic studies found that one-quarter failed tests for specificity, meaning that they often bound to more than one target. Four antibodies were perfectly specific — but to the wrong target.

Scientists often know, anecdotally, that some antibodies in their field are problematic, but it has been difficult to gauge the size of the problem across biology as a whole. Perhaps the largest assessment comes from work published by the Human Protein Atlas, a Swedish consortium that aims to generate antibodies

"ANTIBODIES ARE NOT MAGIC REAGENTS."

for every protein in the human genome. It has looked at some 20,000 commercial antibodies so far and found that less than 50% can be used effectively to look at protein distribution in preserved slices of tissue⁵. This has led some scientists to claim that up to half of all commercially available antibodies are unreliable.

But reliability can depend on the experiment. "Our experience with commercial antibodies is that they are usually okay in some applications, but they might be terrible in others," says Mathias Uhlén at the Royal Institute of Technology in Stockholm, who coordinates the Human Protein Atlas.

Researchers ideally should check that an antibody has been tested for use in particular applications and tissue types, but the quality of information supplied by vendors can vary

tremendously. A common complaint from scientists is that companies do not provide the data required to evaluate a given antibody's specificity or its lot-to-lot variability. Companies might ship a batch of antibodies with characterization information derived from a previous batch. And the data are often derived under ideal conditions that do not reflect typical experiments. Antibody companies contacted for this article said that it is impossible to test their products across all experimental conditions, but they do provide reliable data and work with scientists to improve antibody quality and performance.

Many academics use Google to find products, so optimizing search results can sometimes matter more to a company than optimizing the actual reagents, says Tim Bernard, head of the biotechnology consultancy Pivotal Scientific in Upper Heyford, UK. Christi Bird, a Frost & Sullivan analyst based in Washington DC, says that researchers are often more interested in how quickly reagents can be delivered than in searching for antibodies with appropriate validation data. "It's the Amazon effect: they want it in two or three days, with free shipping."

Researchers who are aware of the antibody problem say that scientists need to be more vigilant. "Antibodies are not magic reagents. You can't just throw them on your sample and expect the result you get is 100% reliable without putting some critical thinking into it," says James Trimmer, head of NeuroMab at the University of California, Davis, which makes antibodies for neuroscience. Like many suppliers, NeuroMab explicitly states the types of experiment that an antibody should be used for, but scientists do not always follow the instructions.

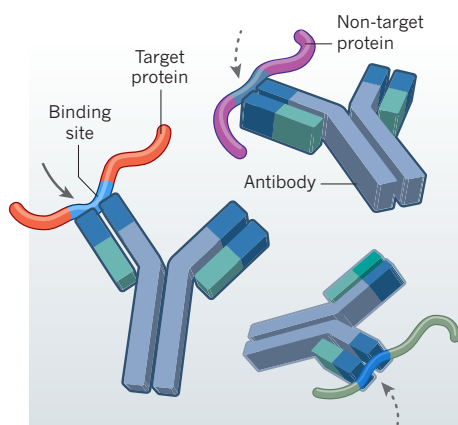
Ideally, researchers would refuse to buy antibodies without extensive validation data or would perform the validation themselves (see 'Bad antibodies'). This is something that Rimm is passionate about: he has developed a multistep flowchart for effective validation⁶, which he shares with anyone who will listen. But the process is time consuming — Rimm recommends control experiments that involve engineering cell lines to both express and stop expressing the protein of interest, for example. Even he acknowledges that few labs will perform all the steps.

Some scientists buy half a dozen antibodies from different vendors, and then run a few assays to see which performs best. But they may end up buying the same antibody from different places. The largest vendors compete on catalogue size, so they often buy antibodies from smaller suppliers, relabel them and offer them for sale. Bernard says that the 2 million antibodies on the market probably represent 250,000–500,000 unique 'core' antibodies.

By necessity, many researchers rely on word of mouth or the published literature for advice. But that creates a self-perpetuating problem, in which better-performing antibodies that become available later are rarely used, says Fridtjof Lund-Johansen, a proteomics

BAD ANTIBODIES

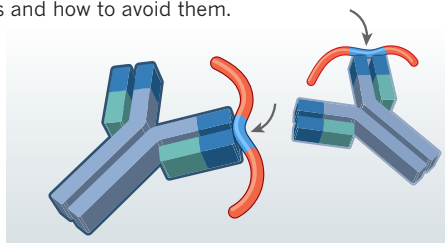
The most common problems with antibodies and how to avoid them.



CROSS-REACTIVITY

Problem: An antibody is supposed to recognize only its target protein, but sometimes binds to others, depending on the proteins present in a sample.

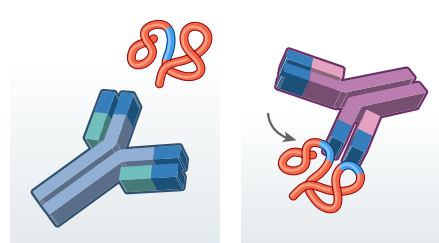
Solution: An antibody should be tested for off-target binding using positive and negative controls.



VARIABILITY

Problem: Separate batches of antibody can perform differently. This happens most often when the antibody is produced from a new set of animals.

Solution: Researchers should confirm lot numbers and characterization data with vendors.



WRONG APPLICATION

Problem: Different experiments and experimental conditions can change a protein's folding and therefore its binding ability.

Solution: Scientists should check supplier's recommended applications.

researcher at the University of Oslo. “We have very good antibodies on the market,” he says, “but we don’t know what they are.” Lund-Johansen is trying to change that by developing high-throughput assays that could compare thousands of antibodies at once.

TESTING TIMES

In the past decade, various projects have sprung up to try to make information about antibodies easier to find. The online reagents portal Antibodypedia (antibodypedia.com), which is maintained by the Human Protein Atlas, has catalogued more than 1.8 million antibodies and rated the validation data available for various experimental techniques. Antibodies-online (antibodies-online.com), another portal, set up a programme two years ago for independent labs to do validation studies, generally at the vendors’ expense. But out of 275 studies, less than half of the products tested have made the cut and earned an ‘independent validation’ badge. The non-profit Antibody Registry (antibodyregistry.org) assigns unique identifiers to antibodies and links them to other resources. Another project, pAbmAbs (pabmabs.com/wordpress), operates in a similar way to the social-recommendation web service Yelp, by encouraging people to review antibodies.

But none of these efforts has gained much of a foothold in the scientific community. Many of the scientists contacted for this article were unaware that such resources existed.

The antibody market has grown so crowded that a reputation for quality is becoming part of some suppliers’ business plans. “Now there is so much competition that you have to differentiate yourself,” says Bernard. Vendors such as Abcam in Cambridge, UK, are encouraging users to report their own data and rankings

on the company’s website. Abcam’s analysis of purchasing behaviour shows that its customers look at data pages on average nine times before buying, suggesting that customers want more information.

Abgent, an antibody company based in San Diego, California, and a subsidiary of WuXi AppTec in Shanghai, China, tested all of its antibodies about a year ago. After reviewing the results it discarded about one-third of its catalogue. Whether that was a good decision depends on whether customers will be willing to spend more for better reagents, says John Mountzouris, site leader at the company. Already, he says, customer complaints have plummeted.

Some scientists are calling for much more radical change. In a Comment in *Nature* in February⁷, Andrew Bradbury of Los Alamos National Laboratory in New Mexico and more than 100 co-signatories proposed a massive shift in the way antibodies are produced and sold. They suggested using only antibodies that have been defined down to the level of the DNA sequence that produces them, and then manufactured in engineered ‘recombinant’ cells. This would circumvent much of the variability introduced by production in animals. But the proposal demands information about individual antibodies that many companies consider to be trade secrets — and the antibody marketplace and its millions of products would have to be essentially demolished and reconstructed.

Uhlén, a co-signatory on the Comment, regards the plan as a distant hope. He estimates that the ‘recombinant antibodies’ that Bradbury hopes for would each cost 10–100 times more to generate than the conventional sort, and that they would not necessarily perform better. “At the end of the day, how the binder works

in the application is more important,” he says. “Having a sequence for sure doesn’t tell you if it works.” Other efforts are under way to find cheap, fast, reliable ways of making antibodies without immunizing animals, for example by expressing and optimizing them in viruses.

The pressure to characterize currently available antibodies is surging. As part of efforts to improve reproducibility, some researchers have started to discuss enlisting an independent body to establish a certification programme for commercial antibodies. And several journals (including *Nature*) ask authors to make clear that antibodies used in their papers have been profiled for that particular application.

The quality will creep, rather than leap, forward, says Trimmer, who hopes to see a positive-feedback loop: as scientists become aware of artefacts, they will be more likely to challenge results and uncover more artefacts. Already, he says, the widespread insouciance about antibody validation has started to fade. “It’s turning around a little bit,” he says. “We need to keep talking about it.” ■

Monya Baker writes and edits for *Nature* in San Francisco, California.

1. Begley, C. G. & Ellis, L. M. *Nature* **483**, 531–533 (2012).
2. Prassas, I. & Diamandis, E. P. *Clin. Chem. Lab. Med.* **52**, 765–766 (2014).
3. Michel, M. C., Wieland, T. & Tsujimoto, G. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **379**, 385–388 (2009).
4. Egelhofer, T. A. et al. *Nature Struct. Mol. Biol.* **18**, 91–93 (2011).
5. Berglund, L. et al. *Mol. Cell. Proteom.* **7**, 2019–2027 (2008).
6. Bordeaux, J. et al. *BioTechniques* **48**, 197–209 (2010).
7. Bradbury, A. & Plückthun, A. *Nature* **518**, 27–29 (2015).



WHAT IS REALLY REAL?

A WAVE OF EXPERIMENTS
IS PROBING THE ROOT OF
QUANTUM WEIRDNESS.

BY ZEEYA MERALI

Owen Maroney worries that physicists have spent the better part of a century engaging in fraud.

Ever since they invented quantum theory in the early 1900s, explains Maroney, who is himself a physicist at the University of Oxford, UK, they have been talking about how strange it is — how it allows particles and atoms to move in many directions at once, for example, or to spin clockwise and anticlockwise simultaneously. But talk is not proof, says Maroney. “If we tell the public that quantum theory is weird, we better go out and test that’s actually true,” he says. “Otherwise we’re not doing science, we’re just explaining some funny squiggles on a blackboard.”

It is this sentiment that has led Maroney and others to develop a new series of experiments to uncover the nature of the wavefunction — the mysterious entity that lies at the heart of quantum weirdness. On paper, the wavefunction is simply a mathematical object that physicists denote with the Greek letter psi (Ψ) — one of Maroney’s funny squiggles — and use to describe a particle’s quantum behaviour. Depending on the experiment, the wavefunction allows them to calculate the probability

of observing an electron at any particular location, or the chances that its spin is oriented up or down. But the mathematics shed no light on what a wavefunction truly is. Is it a physical thing? Or just a calculating tool for handling an observer’s ignorance about the world?

The tests being used to work that out are extremely subtle, and have yet to produce a definitive answer. But researchers are optimistic that a resolution is close. If so, they will finally be able to answer questions that have lingered for decades. Can a particle really be in many places at the same time? Is the Universe continually dividing itself into parallel worlds, each with an alternative version of ourselves? Is there such a thing as an objective reality at all?

“These are the kinds of questions that everybody has asked at some point,” says Alessandro Fedrizzi, a physicist at the University of Queensland in Brisbane, Australia. “What is it that is really real?”

Debates over the nature of reality go back to physicists’ realization in the early days of quantum theory that particles and waves are two sides of the same coin. A classic example is the double-slit experiment, in which individual electrons are fired at a barrier with two

openings: the electron seems to pass through both slits in exactly the same way that a light wave does, creating a banded interference pattern on the other side (see ‘Wave–particle weirdness’). In 1926, the Austrian physicist Erwin Schrödinger invented the wavefunction to describe such behaviour, and devised an equation that allowed physicists to calculate it in any given situation¹. But neither he nor anyone else could say anything about the wavefunction’s nature.

IGNORANCE IS BLISS

From a practical perspective, its nature does not matter. The textbook Copenhagen interpretation of quantum theory, developed in the 1920s mainly by physicists Niels Bohr and Werner Heisenberg, treats the wavefunction as nothing more than a tool for predicting the results of observations, and cautions physicists not to concern themselves with what reality looks like underneath. “You can’t blame most physicists for following this ‘shut up and calculate’ ethos because it has led to tremendous developments in nuclear physics, atomic physics, solid-state physics and particle physics,” says Jean Bricmont, a statistical physicist at the Catholic

DAN HARRIS/MIT



An experiment showing that oil droplets can be propelled across a fluid bath by the waves they generate has prompted physicists to reconsider the idea that something similar allows particles to behave like waves.

Imagine that a cat is enclosed in a steel box. And imagine that the box also contains a sample of radioactive material that has a 50% probability of emitting a decay product in one hour, along with an apparatus that will poison the cat if it detects such a decay. Because radioactive decay is a quantum event, wrote Schrödinger, the rules of quantum theory state that, at the end of the hour, the wavefunction for the box's interior must be an equal mixture of live cat and dead cat.

"Crudely speaking," says Fedrizzi, "in a psi-epistemic model the cat in the box is either alive or it's dead and we just don't know because the box is closed." But most psi-ontic models agree with the Copenhagen interpretation: until an observer opens the box and looks, the cat is both alive and dead.

But this is where the debate gets stuck. Which of quantum theory's many interpretations — if any — is correct? That is a tough question to answer experimentally, because the differences between the models are subtle: to be viable, they

that view the wavefunction as a real entity — psi-ontic models.

To appreciate the difference, consider a thought experiment that Schrödinger described in a 1935 letter to Einstein.

Imagine that a cat is enclosed in a steel box. And imagine that the box also contains a sample of radioactive material that has a 50% probability of emitting a decay product in one hour, along with an apparatus that will poison the cat if it detects such a decay. Because radioactive decay is a quantum event, wrote Schrödinger, the rules of quantum theory state that, at the end of the hour, the wavefunction for the box's interior must be an equal mixture of live cat and dead cat.

"Crudely speaking," says Fedrizzi, "in a psi-epistemic model the cat in the box is either alive or it's dead and we just don't know because the box is closed." But most psi-ontic models agree with the Copenhagen interpretation: until an observer opens the box and looks, the cat is both alive and dead.

But this is where the debate gets stuck. Which of quantum theory's many interpretations — if any — is correct? That is a tough question to answer experimentally, because the differences between the models are subtle: to be viable, they

ace, he says, "there's an overlap and you won't be able to say where it came from". But if you know how many of each type of card is in each deck, you can at least calculate how often such ambiguous situations will arise.

OUT ON A LIMB

A similar ambiguity occurs in quantum systems. It is not always possible for a single measurement in the lab to distinguish how a photon is polarized, for example. "In real life, it's pretty easy to tell west from slightly south of west, but in quantum systems, it's not that simple," says White. According to the standard Copenhagen interpretation, there is no point in asking what the polarization is because the question does not have an answer — or at least, not until another measurement can determine that answer precisely. But according to the wavefunction-as-ignorance models, the question is perfectly meaningful; it is just that the experimenters — like the card-game player — do not have enough information from that one measurement to answer. As with the cards, it is possible to estimate how much ambiguity can be explained by such ignorance, and compare it with the larger amount of ambiguity allowed by standard theory.

That is essentially what Fedrizzi's team tested. The group measured polarization and other features in a beam of photons and found a level of overlap that could not be explained by

"WE WERE TOLD THAT SUCH EFFECTS CANNOT HAPPEN CLASSICALLY, AND HERE WE ARE, SHOWING THAT THEY DO."

University of Louvain in Belgium. "So people say, let's not worry about the big questions."

But some physicists worried anyway. By the 1930s, Albert Einstein had rejected the Copenhagen interpretation — not least because it allowed two particles to entangle their wavefunctions, producing a situation in which measurements on one could instantaneously determine the state of the other even if the particles were separated by vast distances. Rather than accept such "spooky action at a distance", Einstein preferred to believe that the particles' wavefunctions were incomplete. Perhaps, he suggested, the particles have some kind of 'hidden variables' that determine the outcome of the measurement, but that quantum theories do not capture.

Experiments since then have shown that this spooky action at a distance is quite real, which rules out the particular version of hidden variables that Einstein advocated. But that has not stopped other physicists from coming up with interpretations of their own. These interpretations fall into two broad camps. There are those that agree with Einstein that the wavefunction represents our ignorance — what philosophers call psi-epistemic models. And there are those

have to predict essentially the same quantum phenomena as the very successful Copenhagen interpretation. Andrew White, a physicist at the University of Queensland, says that for most of his 20-year career in quantum technologies "the problem was like a giant smooth mountain with no footholds, no way to attack it".

That changed in 2011, with the publication of a theorem about quantum measurements that seemed to rule out the wavefunction-as-ignorance models². On closer inspection, however, the theorem turned out to leave enough wiggle room for them to survive. Nonetheless, it inspired physicists to think seriously about ways to settle the debate by actually testing the reality of the wavefunction. Maroney had already devised an experiment that should work in principle³, and he and others soon found ways to make it work in practice⁴⁻⁶. The experiment was carried out last year by Fedrizzi, White and others⁷.

To illustrate the idea behind the test, imagine two stacks of playing cards. One contains only red cards; the other contains only aces. "You're given a card and asked to identify which deck it came from," says Martin Ringbauer, a physicist also at the University of Queensland. If it is a red

the ignorance models. The results support the alternative view that, if objective reality exists, then the wavefunction is real. "It's really impressive that the team was able to address a profound issue, with what's actually a very simple experiment," says Andrea Alberti, a physicist at the University of Bonn in Germany.

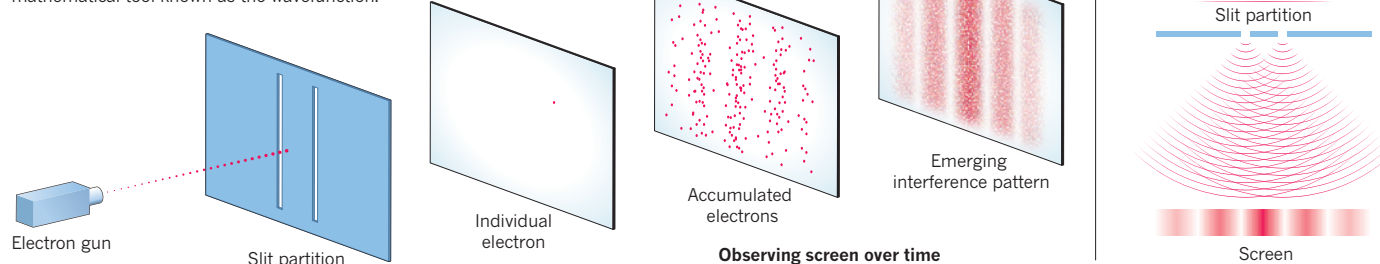
The conclusion is still not ironclad, however: because the detectors picked up only about one-fifth of the photons used in the test, the team had to assume that the lost photons were behaving in the same way⁷. That is a big assumption, and the group is currently working on closing the sampling gap to produce a definitive result. In the meantime, Maroney's team at Oxford is collaborating with a group at the University of New South Wales in Australia, to perform similar tests with ions, which are easier to track than photons. "Within the next six months we could have a watertight version of this experiment," says Maroney.

But even if their efforts succeed and the wavefunction-as-reality models are favoured, those models come in a variety of flavours — and experimenters will still have to pick them apart.

One of the earliest such interpretations

WAVE-PARTICLE WEIRDNESS

When quantum objects such as electrons are fired one by one through a pair of closely spaced slits, they behave like particles: each one hits a screen placed on the far side at exactly one point. But they also behave like waves: successive hits build up a banded interference pattern exactly like that generated by a wave passing through the slits (right). This wave-particle duality is described by a mathematical tool known as the wavefunction.



was set out in the 1920s by French physicist Louis de Broglie⁸, and expanded in the 1950s by US physicist David Bohm^{9,10}. According to de Broglie-Bohm models, particles have definite locations and properties, but are guided by some kind of ‘pilot wave’ that is often identified with the wavefunction. This would explain the double-slit experiment because the pilot wave would be able to travel through both slits and produce an interference pattern on the far side, even though the electron it guided would have to pass through one slit or the other.

In 2005, de Broglie-Bohmian mechanics received an experimental boost from an unexpected source. Physicists Emmanuel Fort, now at the Langevin Institute in Paris, and Yves Couder at the University of Paris Diderot gave the students in an undergraduate laboratory class what they thought would be a fairly straightforward task: build an experiment to see how oil droplets falling into a tray filled with oil would coalesce as the tray was vibrated. Much to everyone’s surprise, ripples began to form around the droplets when the tray hit a certain vibration frequency. “The drops were self-propelled — surfing or walking on their own waves,” says Fort. “This was a dual object we were seeing — a particle driven by a wave.”

Since then, Fort and Couder have shown that such waves can guide these ‘walkers’ through the double-slit experiment as predicted by pilot-wave theory, and can mimic other quantum effects, too¹¹. This does not prove that pilot waves exist in the quantum realm, cautions Fort. But it does show how an atomic-scale pilot wave might work. “We were told that such effects cannot happen classically,” he says, “and here we are, showing that they do.”

Another set of reality-based models, devised in the 1980s, tries to explain the strikingly different properties of small and large objects. “Why electrons and atoms can be in two different places at the same time, but tables, chairs, people and cats can’t,” says Angelo Bassi, a physicist at the University of Trieste, Italy. Known as ‘collapse models’, these theories postulate that the wavefunctions of individual particles are real, but can spontaneously lose their quantum

properties and snap the particle into, say, a single location. The models are set up so that the odds of this happening are infinitesimal for a single particle, so that quantum effects dominate at the atomic scale. But the probability of collapse grows astronomically as particles clump together, so that macroscopic objects lose their quantum features and behave classically.

One way to test this idea is to look for quantum behaviour in larger and larger objects. If standard quantum theory is correct, there is no limit. And physicists have already carried out double-slit interference experiments with large molecules¹². But if collapse models are correct, then quantum effects will not be apparent above a certain mass. Various groups are planning to search for such a cut-off using cold atoms, molecules, metal clusters and nanoparticles. They hope to see results within a decade. “What’s great about all these kinds of experiments is that we’ll be subjecting quantum theory to high-precision tests, where it’s never been tested before,” says Maroney.

PARALLEL WORLDS

One wavefunction-as-reality model is already famous and beloved by science-fiction writers: the many-worlds interpretation developed in the 1950s by Hugh Everett, who was then a graduate student at Princeton University in New Jersey. In the many-worlds picture, the wavefunction governs the evolution of reality so profoundly that whenever a quantum measurement is made, the Universe splits into parallel copies. Open the cat’s box, in other words, and two parallel worlds will branch out — one with a living cat and another containing a corpse.

Distinguishing Everett’s many-worlds interpretation from standard quantum theory is tough because both make exactly the same predictions. But last year, Howard Wiseman at Griffith University in Brisbane and his colleagues proposed a testable multiverse model¹³. Their framework does not contain a wavefunction: particles obey classical rules such as Newton’s laws of motion. The weird effects seen in quantum experiments arise because there is a repulsive force between particles and their clones in parallel universes. “The

repulsive force between them sets up ripples that propagate through all of these parallel worlds,” Wiseman says.

Using computer simulations with as many as 41 interacting worlds, they have shown that this model roughly reproduces a number of quantum effects, including the trajectories of particles in the double-slit experiment¹³. The interference pattern becomes closer to that predicted by standard quantum theory as the number of worlds increases. Because the theory predicts different results depending on the number of universes, says Wiseman, it should be possible to devise ways to check whether his multiverse model is right — meaning that there is no wavefunction, and reality is entirely classical.

Because Wiseman’s model does not need a wavefunction, it will remain viable even if future experiments rule out the ignorance models. Also surviving would be models, such as the Copenhagen interpretation, that maintain there is no objective reality — just measurements.

But then, says White, that is the ultimate challenge. Although no one knows how to do it yet, he says, “what would be really exciting is to devise a test for whether there is in fact any objective reality out there at all.” ■

Zeeya Merali is a freelance writer based in London.

1. Schrödinger, E. *Phys. Rev.* **28**, 1049 (1926).
2. Pusey, M. F., Barrett, J. & Rudolph, T. *Nature Phys.* **8**, 475–478 (2012).
3. Maroney, O. J. E. Preprint at <http://arxiv.org/abs/1207.6906> (2012).
4. Barrett, J., Cavalcanti, E. G., Lal, R. & Maroney, O. J. E. *Phys. Rev. Lett.* **112**, 250403 (2014).
5. Leifer, M. S. *Phys. Rev. Lett.* **112**, 160404 (2014).
6. Branciari, C. *Phys. Rev. Lett.* **113**, 020409 (2014).
7. Ringbauer, M. et al. *Nature Phys.* **11**, 249–254 (2015).
8. de Broglie, L. *J. Phys. Radium* **8**, 225–241 (1927).
9. Bohm, D. *Phys. Rev.* **85**, 166–179 (1952).
10. Bohm, D. *Phys. Rev.* **85**, 180–193 (1952).
11. Couder, Y. & Fort, E. *Phys. Rev. Lett.* **97**, 154101 (2006).
12. Eibenberger, S., Gerlich, S., Arndt, M., Mayor, M. & Tüxen, J. *Phys. Chem. Chem. Phys.* **15**, 14696–14700 (2013).
13. Hall, M. J. W., Deckert, D.-A. & Wiseman, H. M. *Phys. Rev. X* **4**, 041013 (2014).

COMMENT

SUSTAINABILITY Clean-cooking initiatives must help people, not just swap stoves **p.284**

HISTORY When Einstein and philosopher Henri Bergson locked horns **p.286**

PALAEONTOLOGY A bilingual take on vertebrate evolution in China **p.288**



OBITUARY Alexander Rich, discoverer of 'left-handed' DNA, remembered **p.291**

JEROME SESSINI/MAGNUM



Illegal use of opiates such as heroin and morphine affects more than 16 million people worldwide.

Regulate 'home-brew' opiates

The research community and the public require a fast, flexible response to the synthesis of morphine by engineered yeasts, urge **Kenneth Oye, Tania Bubela and J. Chappell H. Lawson.**

Every year, thousands of students from across the world compete to build biological systems from pre-existing parts in a competition organized by the International Genetically Engineered Machine (iGEM) Foundation. Last November, to spark discussion on security and health risks raised by synthetic biology,

FBI Special Agent Edward You presented an example: the production of opiates from sugar by yeast (*Saccharomyces cerevisiae*) that has been genetically modified.

You's hypothetical scenario is becoming a reality. One week after the iGEM competition, two developers of opiate-producing yeast strains approached us, specialists in

biotechnology policy. They had results in advance of publication, and requested advice on how they might maximize the benefits of their research while mitigating the risks. Now, published papers by these researchers — John Dueber at the University of California, Berkeley, and his colleagues¹, and Vincent Martin ▶

► of Concordia University in Montreal, Canada, and his colleagues² — describe all but one step of an engineered yeast pathway that converts glucose to morphine (see ‘Brewing bad’). Meanwhile, researchers at the University of Calgary have put in place the final piece³.

Currently, morphine is produced from the opium poppy (*Papaver somniferum*). By providing a simpler — and more manipulable — means of producing opiates, the yeast research could ultimately lead to cheaper, less addictive, safer and more-effective analgesics. And in generating a drug source that is self-replicating and easy to grow, conceal and distribute, the work could also transform the illicit opiate marketplace to decentralized, localized production. In so doing, it could dramatically increase people’s access to opiates.

In recent years, synthetic biologists have produced numerous benign products — antimalarials, scents, flavours, industrial chemicals and fuels — by modifying yeast, bacteria and eukaryotic plants. Opiate synthesis is the first example of synthetic biology facilitating the production of a controlled narcotic; other new production systems for potentially problematic compounds will almost certainly follow.

The synthetic-biology community, in tandem with regulators, needs to be proactive in evaluating the costs and benefits of such dual-use technologies⁴. Here we lay out the priorities for discussions that are crucial to public health and safety, and to the progress of synthetic biology more broadly. These include restricting engineered yeast

strains to licensed facilities and authorized researchers and technicians; reducing the attractiveness of engineered yeast strains in the illicit marketplace; and implementing a regulatory approach that is flexible and responsive to changes in understanding and capabilities.

COMPLETE PATHWAY

The technology to make morphine from glucose using yeast has been seven years in the making. Three groups of researchers introduced genetic components from poppy, beetroot and a soil bacterium into the yeast genome, creating strains that can perform chunks of the glucose-to-morphine pathway^{1,2,5–7}. A fourth group has developed³ a strain that can convert one of the intermediate compounds, (S)-reticuline, into another, (R)-reticuline. With this final step realized, the creation of a single strain of yeast capable of executing the entire pathway is feasible.

In principle, anyone with access to the yeast strain and basic skills in fermentation would be able to grow morphine-producing yeast using a home-brew kit for beer-making. If the modified yeast strain produced 10 grams of morphine, users would need to drink only 1–2 millilitres of the liquid to obtain a standard prescribed dose. (Current strains are not this efficient, but titres in this range and even

“Yeast-based opiate synthesis could also have an significant effect on illicit markets.”

tenfold higher have been achieved for other commercially relevant metabolic products.)

Although this research is intended to enable synthetic production of opiates for legal pain relief, we perceive several challenges. To be competitive, yeast-based production must be more cost-effective than current systems, more secure and more acceptable to regulators, or provide less addictive, safer products. But most opiates are inexpensive to manufacture, administer and transport.

Advances in breeding high-yield poppies reduced the cost of the main wholesale product, known as concentrate of poppy straw, by 20% between 2001 and 2007 to US\$300–\$500 per kilogram. The design of more commercially valuable opiates will also require collaboration between synthetic biologists, neuroscientists and medicinal chemists among others, and will involve lengthy and expensive clinical trials. What is more, global supply and demand is tightly regulated to limit potential addiction.

LEGAL CONSIDERATIONS

Various international conventions and national laws are designed to prevent diversion to illegal markets. Countries that manufacture opiates commonly use large, secure industrial facilities. Australia further enhances security by growing a thebaine-rich poppy variety; thebaine is toxic to ingest and is not easily converted into morphine. It is difficult to predict how the main international body, the International Narcotics Control Board (INCB), would react to a new production system for opiates. The INCB is unlikely to slash current opium-production quotas and disrupt current legal opiate-trade patterns to accommodate yeast-based production. This would limit the ability of new producers to enter the market.

Meanwhile, yeast-based opiate synthesis could have a significant effect on illicit markets. Currently, opiates are sold illegally through two main channels. First, prescription pain medications such as oxycodone and hydrocodone are pilfered, prescribed improperly or prescribed legitimately but then sold on illegally by patients. Second, illegally cultivated opium poppies in countries such as Afghanistan, Myanmar, Laos and Mexico are processed into heroin and distributed by criminal networks that sell them at street prices several dozen times the production costs⁸.

Yeast-based production of opiates could provide an alternative system for current criminal networks, particularly in North America and Europe, where the drugs are in high demand. Because yeast is easy to conceal, grow and transport, criminal syndicates and law-enforcement agencies would

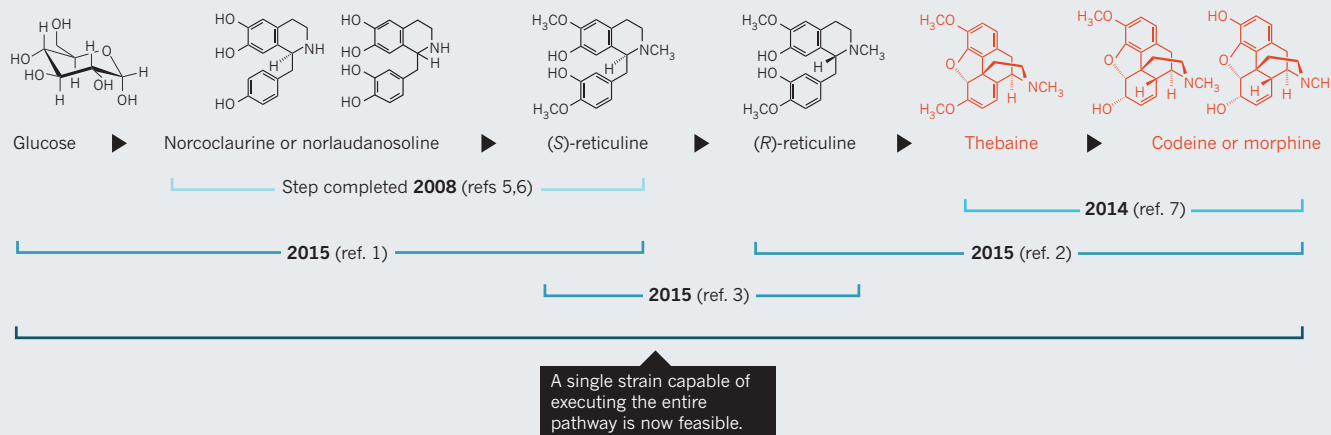


In principle, a home-brew kit for beer-making could be used to make morphine.

HOW HWEE YOUNG/EPA/CORBIS

BREWING BAD

Researchers have completed all the steps of an engineered pathway in yeast that make the controlled substances thebaine and morphine from glucose.



have difficulty controlling the distribution of an opiate-producing yeast strain. All told, decentralized and localized production would almost certainly reduce the cost and increase the availability of illegal opiates — substantially worsening a worldwide problem. Globally, more than 16 million people use opiates illegally.

FOUR RECOMMENDATIONS

There are two major challenges to developing and implementing a flexible and proportionate regulatory approach for this research. Current regulations for engineered organisms focus on pathogenic organisms such as the anthrax bacterium and smallpox, not on yeasts. And the array of national and international drug regulators and law-enforcement agencies that would need to be involved have different practices and norms.

Increased communication and coordination will be required among public-health experts, scientists, regulators and law-enforcement agencies. Potential international focal points for dialogue are the INCB and the international expert groups on biosafety and biosecurity regulation.

The following four issues warrant immediate consideration.

Engineering. Yeast strains should be designed to make them less appealing to criminals. For example, strains could be engineered to make only opiates with limited street value, such as thebaine. Alternatively, weaker strains could be engineered to make it harder for people to cultivate and harvest opiates outside established laboratory settings. Strains could be engineered with unusual nutrient dependencies, for instance. Such methods of ‘biocontainment’ have been developed in *Escherichia coli*. Opiate-producing yeast strains could also

contain a marker, such as a DNA watermark, that makes them more readily identifiable to law-enforcement agencies.

Screening. Because there is some — albeit low — risk of criminal syndicates synthesizing opiate-producing yeast strains using published DNA sequences, commercial organizations that make stretches of DNA to order should be alerted. The sequences for opiate-producing yeast strains should be added to the screening criteria used by these providers. Overseen by two voluntary consortia, the International Association of Synthetic Biology and the International Gene Synthesis Consortium, these criteria currently cover only pathogens.

Security. Efforts should be made to keep opiate-producing yeast strains in controlled environments that are licensed by regulators. Physical biosecurity measures — including locks, alarms and systems for monitoring the use of laboratories and materials — could help to prevent the theft of yeast samples. Laboratory personnel should be subject to security screening. Similarly, assigning liability and penalties may dissuade researchers from sharing strains with anyone who is not legally authorized to work with them.

Regulation. The current laws covering opiates, such as the US Controlled Substance Act and its worldwide equivalents, should be extended to cover opiate-producing yeast strains, to make their release and distribution illegal.

The right choices in the regulation of this dual-use technology will set a precedent for other fast-emerging biotechnologies. In fact, biologists working on yeast-based

opiates have already led the way on the most important aspect — namely, their willingness to take responsibility for the tools they are developing. But for them, this article would not have been written.

Other genomic engineers are following this path. Developers of the gene-editing tool CRISPR/Cas9 have called for proactive engagement with risks before altering populations of animals and plants in the wild or manipulating human reproductive cells^{9,10}. With all the signs that synthetic biology is coming of age, this type of responsible conduct is imperative. ■ **SEE NEWS P.267**

Kenneth A. Oye and J. Chappell H. Lawson are associate professors in the Department of Political Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. **Tania Bubela** is professor and associate dean for research at the School of Public Health, University of Alberta, Edmonton, Canada. e-mail: oye@mit.edu

- DeLoache, W. C. et al. *Nature Chem. Biol.* <http://dx.doi.org/10.1038/nchembio.1816> (2015).
- Fossati, E., Narcross, L., Ekins, A., Falgoutyret, J. P. & Martin, V. J. J. *PLoS ONE* **10**, e0124459 (2015).
- Beaudoin, G. A. W. *Characterization of Oxidative Enzymes Involved in the Biosynthesis of Benzylisoquinoline Alkaloids in Opium Poppy (Papaver somniferum)*. PhD thesis, Univ. Calgary (2015); available at <http://hdl.handle.net/11023/2115>
- Bubela, T., Hagen, G. & Einsiedel, E. *Trends Biotechnol.* **30**, 132–137 (2012).
- Hawkins, K. M. & Smolke, C. D. *Nature Chem. Biol.* **4**, 564–573 (2008).
- Fossati, E. et al. *Nature Commun.* **5**, 3283 (2014).
- Thodey, K., Galanie, S. & Smolke, C. D. *Nature Chem. Biol.* **10**, 837–844 (2014).
- International Narcotics Control Board. *Report of the International Narcotics Control Board on the Availability of Internationally Controlled Drugs: Ensuring Adequate Access for Medical and Scientific Purposes* (INCB, 2011).
- Oye, K. A. et al. *Science* **345**, 626–628 (2014).
- Baltimore, D. et al. *Science* **348**, 36–38 (2015).



Women in Guatemala cook with a solar oven that they built themselves.

Clean cooking empowers women

Putting women and girls at the centre of solar-oven programmes builds communities and reduces pollution, say **Laura S. Brown** and **William F. Lankford**.

At 4 a.m., Eunice in El Jobo, Nicaragua, puts a fistful of *madero de madura* (seasoned firewood) in her mud and brick oven. The wood catches flame; a cloud of smoke and ash rises, resting beneath the metal roof. Eunice launches into a coughing fit. Her daily routine has started.

Around the world, 3 billion poor people

like Eunice cook over unventilated wood, coal or dung fires. About 4 million deaths a year are associated with inhaling the smoke¹. Hours spent gathering wood or biomass fuel robs families of time and energy needed for education and work. Deforestation degrades soil and destroys habitats.

Massive efforts are under way to introduce

cleaner ways of cooking, such as with solar ovens. Most of these attempts are too narrowly focused. Stoves are replaced without channelling women's time savings and health improvements into community development.

Women who must cook with solid fuels face many other obstacles, such as poor health and malnutrition, lack of transportation, illiteracy, ingrained social attitudes about their roles and potential, and resistance to change from beneficiaries of the status quo.

More-holistic approaches can address a constellation of difficulties, as shown by our long experience with the Central American Solar Energy Project (CASEP, of which W.F.L. is president and L.S.B. is programme coordinator). CASEP is a private foundation funded mainly through family donations. For nearly 25 years, it has provided financial and technical assistance to thousands of poor, rural women for the construction of solar ovens in Guatemala, Costa Rica, Nicaragua and Honduras. CASEP weaves into its programmes opportunities for women to break gender barriers and receive education on health, human rights, leadership and community engagement.

Our modest programme is a worked example of the kind of on-the-ground change that is needed more broadly at the nexus of international development, clean-energy markets and sustainable technologies. We call on decision-makers in these fields to reframe their work around supporting women to adopt clean cooking by following four key practices (see 'Lifting lives, not just swapping stoves'). These steps encourage individual, household and community development while reducing carbon emissions.

WIDER BENEFITS

Solar-cooking initiatives face challenges. Critics emphasize the difficulty of developing financially sustainable markets for solar ovens. The most durable stoves are expensive and difficult to manoeuvre or repair. Cheap ovens are insubstantial and easily stolen. Solar cookers that are designed to focus the Sun's rays with parabolic mirrors are vulnerable to wind damage and require constant repositioning to catch sunlight. Opportunities for generating profit are limited — no one sells sunshine.

Distributing subsidized solar ovens also has drawbacks. International-development programmes are results-driven and focus on easily measurable impacts. They do not draw connections between women's cooking needs and their other needs. The focus is the stove, not the people who use it.

The most prominent clean-cooking effort is the Global Alliance for Clean Cookstoves. Launched five years ago by then-US secretary of state Hillary Clinton, the alliance counts 1,000 government, non-profit, academic and business groups as partners. Its goal is

to convert 100 million households to clean cooking fuels by 2020 and to establish a thriving market for clean stoves, with a focus on eight countries: Nigeria, Kenya, Uganda, Ghana, Guatemala, Bangladesh, India and China. Solar power is not a focus: of 278 stove designs promoted, only 13 are sun-powered.

CASEP has a different approach. Zero greenhouse-gas emissions and free fuel make solar ovens the cleanest and most accessible option for many poor households, costing around US\$300 each in materials and workshop costs (see go.nature.com/pvweju).

When CASEP was launched in 1991, local tradesmen built the ovens and presented them to their wives. The women took the gesture as criticism — what was wrong with how they had cooked all those years? The ovens gathered dust. So we put women at the heart of the programme.

KEEPING MOMENTUM

Today, at solar-cooking demonstrations, we identify 10–20 women who wish to build an oven. We train some as instructors and provide materials. The women invest a lot of work and time, spending up to three full-time weeks or ten weekends working on the stoves.

Women construct them from plywood, glass, aluminium and metal. They use hammers, hand saws, shears and sealant guns — items they had considered ‘men’s tools’. Participants complete each step of the construction process by rotating among several work stations. Building the ovens collectively and in stages nurtures a spirit of solidarity.

The workshop serves as a proving ground for leadership skills. Participants elect a leadership team to manage the workshop finances, meals and interpersonal issues. Once the ovens are complete, graduates celebrate in public at the *clausura*, a solar banquet with music and festivities. “This oven that we women have made with our very own hands is

a source of pride for me ... it’s a great advantage,” said Maria Francisca, a workshop graduate in Nicaragua. Adequate follow-up is key

to the adoption of solar cooking. Ongoing support is provided for at least two years by CASEP staff in each country.

Women who use solar ovens tell us that they save the time spent on gathering firewood and the money spent on fuel; are able to leave their meals to cook unattended; and enjoy clean indoor air and improved health. And the benefits go further.

Women apply their new skills to advocate for themselves and others. They form solidarity groups, feeding programmes and micro-loan enterprises. They develop

“The focus of international-development programmes is the stove, not the people who use it.”

FOUR STEPS TO CLEAN COOKING

Lifting lives, not just swapping stoves

1 Develop flexible, integrated, long-term approaches.

Poor women and communities face multiple difficulties; it is not possible to fix only one component.

The women should be asked: what strengths and challenges do they see in their communities? How would they like to respond? What do they need to realize that response? Monthly visits, consultations and repair assistance should be funded for several years after ovens are introduced³.

2 Promote solar ovens worldwide as companions to other clean stoves.

Women often persist with solar cooking until their income rises. Then, they opt for the methods of the middle class: stoves powered by natural gas or electricity. Policy-makers should advance and incentivize solar cooking as the cleanest option. Research is needed to establish the effectiveness of solar ovens used in conjunction with improved stoves such as those that use liquified petroleum gas or compressed biomass pellets in reducing carbon emissions and greenhouse gases.

Quality standards that better quantify ‘clean fuels’ should recognize solar cooking as the gold standard.

3 Focus research on the nexus of women, energy and poverty.

More social-science studies are needed to probe the ways in which poor women receive support during the conversion away from their conventional methods of cooking. Researchers should investigate how it affects daily routines, what shapes the ways they invest their new-found time and what cultural factors hinder the conversion to clean cooking fuels.

4 Evaluate the benefits for future generations.

Researchers should quantify the effects of cooking with clean stoves on all members of the household, particularly young people. Seeing three girls cooking with their mother’s solar oven in Guatemala in 2010 demonstrated to us at CASEP that for the next generation, cooking with solar ovens could be as natural as cooking with wood.

ecological household systems for water and waste, organic community farms and urban agriculture programmes. “They emerge as new women: women without limits, who value themselves, who have rights, who refuse to be limited or violated by anyone,” said Fatima, a participant in Costa Rica².

CASEP’s affiliate group in Honduras, the Association of Women Defenders of Life, manages a micro-loan programme, leads efforts to prevent maternal and child malnutrition and facilitates leadership and civic-engagement programmes for women and young people.

In Costa Rica, CASEP gave rise to Casa del Sol (House of the Sun), a demonstration centre for household solar applications such as cooking, water purification and lighting. And Sol de Vida (Sun of Life), which constructs solar ovens, works for gender equality and engages in ecological activism.

In Nicaragua, the Solar Project Foundation for Nicaraguan Women won the 2013 Energy Globe World Award in the air category, for its construction of solar ovens in rural and semi-urban communities.

There have been setbacks. Some women’s associations have stopped. Ovens have been abandoned. Programmes have ceased owing to corruption and lack of leadership. Finding and retaining skilled female leaders is CASEP’s biggest challenge. Many of the programme’s beneficiaries become local

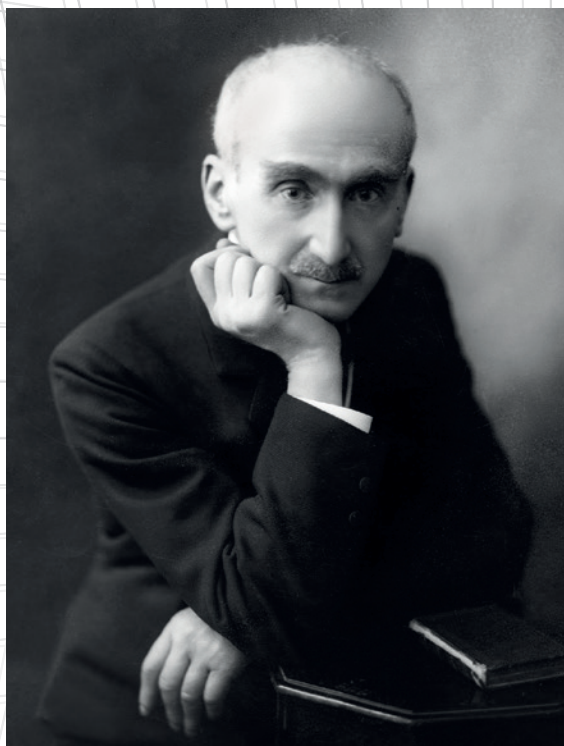
community leaders and advocates; few want to develop the professional skills needed to lead an international organization.

The difficulty of adapting lifestyles to cook with sunlight is also a challenge. Although a CASEP solar oven is simple to repair, long-lasting and large enough to cook a full meal for a typical family, it requires longer cooking times that depend on the weather. Households that use solar cooking need to change how and when they prepare food, and need ongoing support.

Leaders in public policy, sustainable technologies, emerging energy markets and international development should focus on encouraging women to adopt clean cooking practices. Conversion efforts towards cleaner stoves that leverage savings in women’s time and resources for the common good will amplify the benefits and create community change. ■

Laura S. Brown is programme coordinator and **William F. Lankford** is founder and president of the Central American Solar Energy Project (CASEP), Charlottesville, Virginia, USA.
e-mail: casep12@gmail.com

1. Subramanian, M. *Nature* **509**, 548–551 (2014).
2. Vega, N. M. *Sol de Vida y La Red Solar* 5–6 (Univ. Costa Rica, 2001).
3. Kammen, D. M. & Lankford, W. F. *Nature* **348**, 385–386 (1990).



Henri Bergson (left) thought Albert Einstein's theory of relativity was a flawed philosophy.

PHYSICS

Fighting for time

Graham Farmelo enjoys an account of Einstein's clash with philosopher Henri Bergson.

Bertrand Russell is reputed to have said that “science is organized common sense; philosophy is organized piffle”. Although probably being playful, he was articulating the view of many physicists. Theoretical physicist Steven Weinberg declared the “unreasonable ineffectiveness” of philosophy in his field; he was outdone by Stephen Hawking, who in 2011 pronounced philosophy “dead”. Yet only a century ago, the two disciplines coexisted happily.

One theoretician who read widely in philosophy was Albert Einstein. Physicist Nándor Balázs, who worked with him in the early 1950s, told me that Einstein would often spend hours reading philosophy, and admired the work of seventeenth-century Dutch philosopher Baruch Spinoza. However, he had little time for those who expatiated on physics that they did not understand. This seems to have been at the root of tensions between Einstein and French philosopher Henri Bergson. Their quarrel about the nature of time is the subject of *The Scientist and the Philosopher*, a hefty, stimulating study by science historian Jimena Canales.

Canales begins with an account of their meeting, at the French Philosophical Society

The Physicist and the Philosopher: Einstein, Bergson, and the Debate That Changed Our Understanding of Time

JIMENA CANALES

Princeton Univ. Press: 2015.

in Paris on 6 April 1922. Bergson was 62 and had long been internationally famous. Einstein, two decades his junior, had recently become an even more prominent celebrity, after astronomers gave widely publicized empirical support to his general theory of relativity.

Their exchange was intellectually sterile. We do not know exactly what Bergson said, but he probably expressed the views set out in his contentious *Duration and Simultaneity* later that year. In it, he chastised relativity theory for going beyond physics to become a “flawed philosophy” that should be strongly resisted. He felt that human consciousness plays a crucial part in our knowledge of the Universe, so a complete account of time must reflect its subjective aspects (our perception of durations of time depend, of course, on the circumstances in which we experience them).

Bergson spent half an hour putting his case; it was certain to raise the hackles of Einstein, who strove to remove subjective elements

from his theories. Einstein's reply was terse to the point of rudeness. He said that there were only two ways of understanding time — psychological and physical — and the philosopher's time did not exist. The rebuttal lasted about a minute. That night, Einstein wrote to his wife: “All went brilliantly well.” He believed that Bergson was confused and ignorant about relativity. Bergson was convinced that his opponent had not understood him.

Bergson plainly did not comprehend basic aspects of relativity, so it is hardly surprising that this spat did nothing to make leading theoreticians reassess the theory. But he did some damage. In 1922, when Einstein received a Nobel Prize for his “services to theoretical physics”, the citation mentioned his work on the photoelectric effect, not relativity. Pressed to explain, Nobel Committee president Svante Arrhenius said: “It will be no secret that the famous philosopher Bergson in Paris challenged the theory.” Four years later, Bergson was awarded a Nobel of his own, for literature.

Canales aims to clarify the essence of the quarrel without taking sides. Reading between the lines, she seems to sympathize with maverick twentieth-century physicist

L: EINSTEIN BILD VIA GETTY; R: MARTIN HOHLGUT/EINSTEIN BILD VIA GETTY

and critic of relativity theory Herbert Dingle, who lamented that in general the scientist “understands what he is doing about as well as a centipede understands how he walks”.

Einstein does not seem to have spent much time worrying about Bergson's views, although he commented on the meeting occasionally to friends, not giving any ground. Bergson, by contrast, criticized Einstein's relativistic concept of time and promoted his own case indefatigably. He wrote to Einstein's hero, physicist Hendrik Lorentz, who despite differences with Einstein offered little solace. Bergson also conversed with Albert Michelson — a top-drawer experimentalist but not a deep thinker about relativity — and used his insights to inform his case against Einstein. Canales does sterling work investigating these engagements, and even the largely incoherent contributions of the Catholic Church. Regardless of the views of his few critics, Einstein's concept of time, supported by experiment, became part of the bedrock of physics.

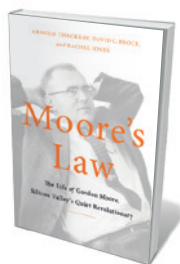
In my view, Canales exaggerates Bergson's influence on our understanding of time, and underestimates Einstein's substantial contribution to philosophy. Throughout his career, he was thoughtful about the philosophy of physics. With colleague Max Planck, he even helped to create a chair in the philosophy of science at the University of Berlin in the mid-1920s. Around that time, the value of philosophy was discounted by several young pioneers of quantum mechanics, notably Paul Dirac.

Canales oddly portrays the development of quantum physics as an embarrassment for Einstein, when he was not only one of its pioneers but also perhaps its most astute and respected critic. She sees the theory as a vindication of Bergson, whom she credits with anticipating Werner Heisenberg's uncertainty principle by some 20 years. It seems to me fanciful to link Bergson's long advocacy of indeterminism with Heisenberg's precise concept of the indeterminability of specific pairs of variables in quantum mechanics. Nor does Canales underline that physicists produced a fully relativistic theory of quantum mechanics, incorporating Einsteinian time.

She does a fine job, however, of highlighting the lack of constructive engagement between physicists and philosophers, beyond a few centres that specialize in the philosophy of physics. I sense that many would like to see some sort of rapprochement, and I warmly agree. On the evidence presented in this stimulating book, however, such a revolution is likely only after physicists shed some of the condescension that they sometimes show to other disciplines, and after philosophers cut from their discourse every last trace of piffle. ■

Graham Farmelo is a by-fellow at Churchill College of the University of Cambridge, UK, and author of Churchill's Bomb.
e-mail: graham@grahamfarmelo.com

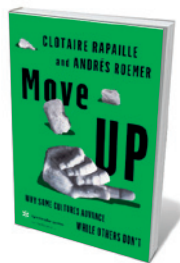
Books in brief



Moore's Law: The Life of Gordon Moore, Silicon Valley's Quiet Revolutionary

Arnold Thackray, David C. Brock and Rachel Jones BASIC (2015)

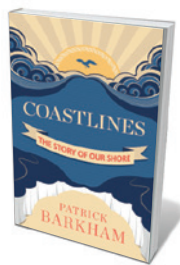
In 1957, experimental chemist Gordon Moore and his colleagues formed a start-up manufacturing silicon transistors in Mountain View, California. Silicon Valley was born, and the prediction known as Moore's Law began to play out: the number of transistors in integrated circuits started to double every two years. Arnold Thackray, David Brock and Rachel Jones transform Moore from a man “doing something inscrutable in the margins” to a comprehensible, fiercely driven technophile who shaped history from the inside out.



Move UP: Why Some Cultures Advance While Others Don't

Clotaire Rapaille and Andrés Roemer ALLEN LANE (2015)

With gross domestic product looking ever thinner as an index of success, marketing specialist Clotaire Rapaille and diplomat Andrés Roemer proffer a new analytic tool for gauging progress, informed by behavioural economics, neuroscience and evolutionary psychology. Their R² Mobility Index rests on a country's cultural capacity to enable upward mobility, and its ability to sensibly support the basic biological imperatives of security, success, survival and sex. Scandinavian nations top several indices here, but Rapaille and Roemer's provocative synthesis throws up surprises too.



Coastlines: The Story of Our Shore

Patrick Barkham GRANTA (2015)

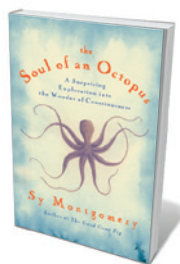
“The British Isles,” writes Patrick Barkham, “are more edge than middle.” Here he pays homage to the chalk cliffs and tidal flats of the 17,800-kilometre coastline to mark 50 years of National Trust protection of more than half of it. Filtered through his hyper-observant sensibility, it all becomes fabulously strange: Undercliff near Lyme Regis, for instance, is an active landslide festooned with botanical oddities and criss-crossed by shrews. Barkham's tour of the wind-scoured spots on this ragged borderland reminds why it has mesmerized scientists, artists and all those hungering for horizons.



Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up

Philip N. Howard YALE UNIVERSITY PRESS (2015)

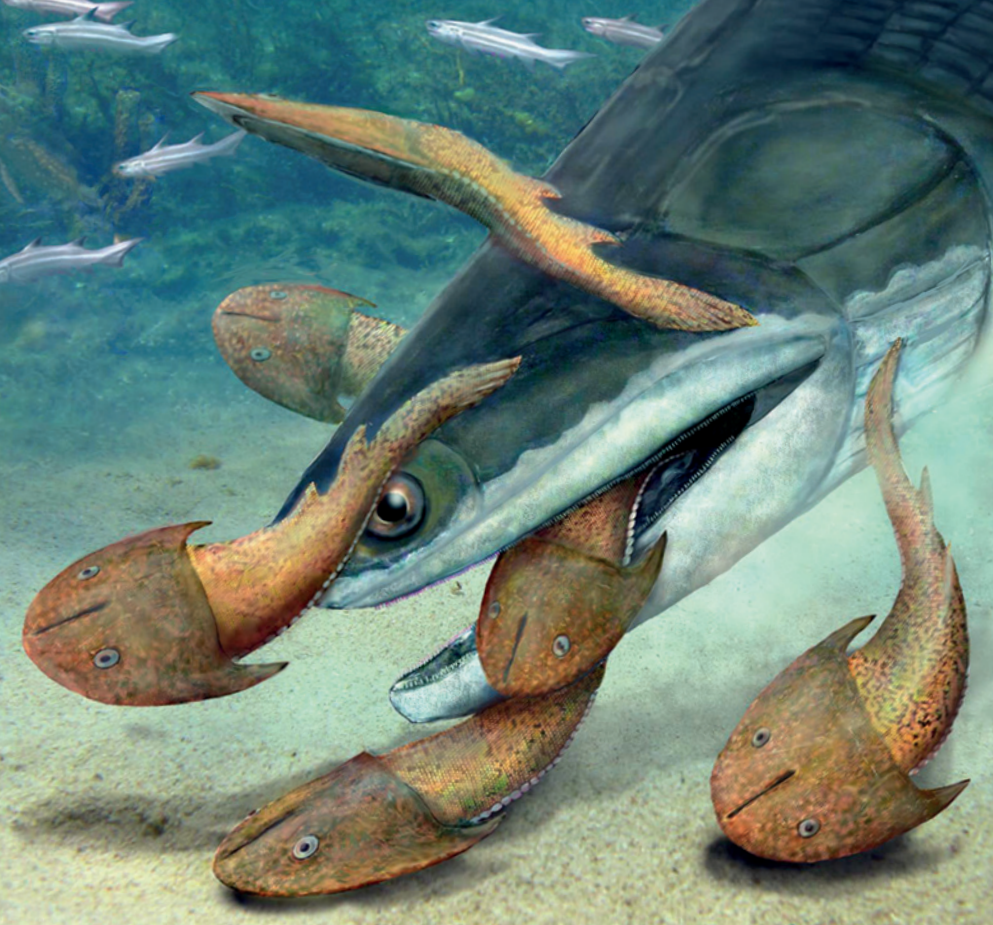
The Internet of Things could encompass 30 billion connected smart devices — from cars to spectacles — within just five years. In analysing this pervasive phenomenon, sociologist Philip Howard emphasizes its potential as the titular “pax technica”, binding industry and government in “mutual defense pacts, design collaborations, standards setting and data mining”. Howard duly notes possible risks, such as intensified mass surveillance, but argues that new devices could become “liberation technologies”.



The Soul of an Octopus: A Surprising Exploration into the Wonder of Consciousness

Sy Montgomery ATRIA (2015)

“Twisting, gelatinous, her arms boil up from the water, reaching for mine.” So begins naturalist Sy Montgomery's close encounter with a giant Pacific octopus (*Enteroctopus dofleini*) in this delightful study of cephalopods in the wild, aquaria and labs. Montgomery celebrates the solitary invertebrates in all their behavioural and physiological glory — as playful escapologists, problem-solvers and masters of camouflage that can taste and might even see with their skin. **Barbara Kiser**



Megamastax amblyodus eating *Dunyu longiferus*, from the 423-million-year-old Xiaoxiang Biota of China.

From Fish to Human: The March of Vertebrate Life in China

CORWIN SULLIVAN,
WANG YUAN AND
BRIAN CHOO
Popular Science:
2015.

Nature <http://doi.org/4j4; 2013>).

Each fauna or group of faunas is followed by a concise review of a related evolutionary transition. So the discussion of how feathers evolved into their modern form fol-

lows a description of the Jehol Biota, which yielded much of the fossil evidence for it.

The balance between specific discoveries and general evolutionary history allows a clear and current understanding of vertebrate evolution, and showcases the beauty of the extinct animals. Thus stunning photographs, for example of the exquisitely preserved fossils from the Chengjiang Biota, sit alongside remarkable reconstructions of creatures and habitats. Informative drawings show the family trees of major vertebrate groups and biological structures such as tetrapod limbs.

I do have quibbles. The chapter on basics such as how fossils form will help beginners, but the rest seems to be written for advanced readers. Discussion of some important groups is missing; for example, we read nothing of the Late Cretaceous Bayan Mandahu Fauna (about 75 million years old), from which significant information on dinosaur behaviour has been recovered. The authors list 79 major vertebrate fossil sites in China — at least 10 short, in my view. They describe only 15.

Attractive global palaeogeographic maps show the arrangement of the continents and oceans in different geological periods, but there is no indication of the modern locations of the fossil faunas in China. Neither is there a summary guide to their stratigraphic distribution. And in a few cases, the Chinese translation is subtly different in meaning and intent from the original English.

Nevertheless, this book should occupy the shelf of anyone eager to keep up with advances in palaeontology and evolution, or to know more about Chinese palaeontology. ■

Xu Xing is a professor at the Institute of Vertebrate Paleontology and Paleoanthropology of the Chinese Academy of Sciences in Beijing.
e-mail: xuxing@ivpp.ac.cn

CORRECTION

The Q&A 'Geological historian' (*Nature* **520**, 294; 2015) incorrectly used "geology and surveying" instead of "geometry and surveying", and "core seams" instead of "coal seams".

BRIAN CHOO

PALAEONTOLOGY

Tracing the backbone in China's rocks

Xu Xing relishes a bilingual book on the evolution of vertebrate life in his fabulously fossil-rich country.

China's rich fossil resources have supplied many firsts — discoveries that have rewritten and helped to construct evolutionary history. The bilingual (English and Chinese) *From Fish to Human* summarizes and highlights the spectacular Chinese vertebrate fossil record and its place in the broader span of vertebrate life.

This volume, rich with illustrations, was produced by an international team of vertebrate palaeontologists. Corwin Sullivan wrote the English text with input from the other authors, Wang Yuan did the Chinese translation and Brian Choo produced the illustrations. (I work alongside all three, but was not involved with this book.) Their effort has produced an excellent resource.

From Fish to Human describes 15 Chinese faunas — collections of fossils of a similar age from the same general area — that highlight every major geological time period from the early Cambrian to the late Pleistocene

epoch. The Cambrian Chengjiang Biota, for example, dating to around 525 million years ago, contains the oldest known diverse multicellular animals, including the earliest vertebrates, such as the primitive, fishlike *Haikouichthys*. It is shedding light on the rapid diversification of life known as the Cambrian explosion.

Flamboyant feathered dinosaurs from the roughly 125-million-year-old Jehol Biota, such as the four-winged *Microraptor* and the gigantic tyrannosaur *Yutyrannus*, have garnered much public attention; less known are the earlier Silurian Xiaoxiang and Devonian Zhongning faunas. The authors explain how discoveries from these have helped to establish the evolution of important structures such as jaws and limbs. *Entelognathus* from Xiaoxiang, for example, is a placoderm, or armoured fish, with the jawed face of an osteichthyan, or bony fish — a finding that blurs the boundary between major vertebrate groups (see

Correspondence

Engage public in gene-editing policy

I agree that new mutational technologies such as gene editing and gain-of-function research call for public debate, global engagement and broad evaluation by experts so that policy-makers are properly informed (see *Nature* **521**, 5; 2015).

The degree of experimental freedom in research is not the only thing at stake, however. At issue is public trust in the institution of science itself. Similar policy challenges arose when recombinant DNA technology was first developed, which led to safety guidelines being drawn up at the 1975 US Asilomar conference.

An international set of such conferences is now needed to assess the potential risks associated with the latest DNA technologies and to develop a common understanding of where lines should be drawn (see also E. Lanphier *et al.* *Nature* **519**, 410–411 (2015) and D. Baltimore *et al.* *Science* **348**, 36–38 (2015)).

The original Asilomar meeting failed to engage the public in discussions, which we now know is crucial to the regulatory decision-making process. Had it done so, the resulting guidelines on recombinant DNA might have extended to legislation covering all users — including the military and commercial sectors — and not just those funded by the US National Institutes of Health.

Filippa Lentzos *King's College London, UK.*
filippa.lentzos@kcl.ac.uk

Blood-transfusion decisions not simple

We consider that your discussion on the possible overuse of blood transfusions simplifies a complex issue (*Nature* **520**, 24–26; 2015).

Readers might infer, for example, that a standardized transfusion protocol is safer than individualized blood-management care, or that

restricted transfusion in response to a particular haemoglobin concentration is at least as safe and effective as transfusion titrations determined by a range of haemoglobin levels. In fact, neither hypothesis has been tested. As you indicate, physicians need to take into account individuals' primary disease as well as any complications and accompanying disorders. Such factors can improve patient care in the longer term, beyond any simple numerical guidelines for restrictive-transfusion practices.

We have shown that a restrictive-transfusion threshold at 7 grams per decilitre of haemoglobin can be problematic for people with stable blood pressure and cardiovascular status (H. G. Klein and C. Natanson *Ann. Intern. Med.* **157**, 753–754; 2012). We argue elsewhere that the pivotal trials you cite would have been more robust had they included a range of transfusion triggers, instead of just two arbitrarily selected haemoglobin levels, and a standard-of-care control arm to incorporate all clinical and lab observations for each patient (K. J. Deans *et al.* *Vox Sang.* **99**, 16–23; 2010). Restrictive practices save blood, but they do not necessarily save lives.

Harvey G. Klein, Irene Cortés-Puch, Charles Natanson *National Institutes of Health, Bethesda, Maryland, USA.*
hklein@dtm.cc.nih.gov

Water: megacities running dry in Brazil

São Paulo and Rio de Janeiro are running out of drinking water owing to an extended drought and disjointed water-resource planning in Brazil. To avert social, economic and political disruption, scientific information must be translated more effectively into water policy.

For example, industrial sectors need information on how to adapt hardware, methods and

practices to mitigate the water shortage; state and regional governments need guidance on economic-development models that are tailored to the capacities of regional freshwater ecosystems.

Educating the public and all stakeholders in water usage is a priority, and is a primary function of Brazil's International Centre for Education, Capacity Building and Applied Research in Water (HidroEX). Political organizations must promote responsible water use, as is proving successful in California.

Richard Meganck *Oregon State University, Corvallis, USA.*
Karl Havens *University of Florida, Gainesville, USA.*
Ricardo M. Pinto-Coelho *Federal University of Minas Gerais, Brazil.*
rameganck@gmail.com

Water: halt India's groundwater loss

India is not doing enough to stop groundwater depletion (see M. Rodell *et al.* *Nature* **460**, 999–1002; 2009 and P. P. Mujumdar *Nature* **521**, 151–155; 2015). This could compromise its capacity to resolve food-security issues in the face of climate change.

India's water storage per head of population is only 200 cubic metres, compared with around 2,500 m³ in China and almost 6,000 m³ in the United States. Farmers in India are digging ever deeper for water, unaware of the long-term repercussions.

Government plans for more-efficient water usage include the National Water Mission and National Water Policy. Water regulation, management and monitoring are regionally but not federally controlled, so enforcing such policy initiatives in different states is likely to take several years.

A national water project to irrigate some 35 million hectares of land through river linking and extensive canal networks could prove impractical because of its cost (US\$92 billion, or around 5% of India's 2013 gross domestic product) and because of the

political complexities associated with rivers that cross other countries.

None of these plans will reduce groundwater depletion in the short term. They need to be supported with public education on water use and with integrated water-resource management. Leading local policy-makers and central administration must cooperate to meet the challenge before it is too late.

Bobban Subhadra *Quorum Innovations, Sarasota, Florida, USA.*
bbobban@gmail.com

Water: a drought plan for biodiversity

To help combat California's worst drought for more than 1,000 years, state governor Jerry Brown has called for the replacement of urban lawns with drought-tolerant landscaping (see go.nature.com/cvqw4l). Applied on a larger scale than he proposes, this move would boost the region's threatened biodiversity and improve water conservation.

The governor's designated "50 million square feet of lawns" amounts to 4.64 square kilometres, or just 0.04% of the estimated 11,000 km² of turf grass in California (C. Milesi *et al.* *Environ. Manage.* **36**, 426–438; 2005). Planting drought-resistant native vegetation instead of non-native turf grasses would spare water and help to restore native ecosystems in the longer term.

Ecologists should work with municipal initiatives to expand such refurbished urban green spaces into nearby areas of native vegetation, which would restore connectivity between native-habitat remnants. This model could be applied in other drought-stricken regions with high biodiversity, such as those in Australia and southeast Brazil.

Alexander C. Lees *Museu Paraense Emílio Goeldi, Belém, Pará, Brazil.*
Peter Bowler *University of California, Irvine, USA.*
alexanderlees@btopenworld.com

Alexander Rich

(1924–2015)

Biologist who discovered ribosome clusters and ‘left-handed’ DNA.

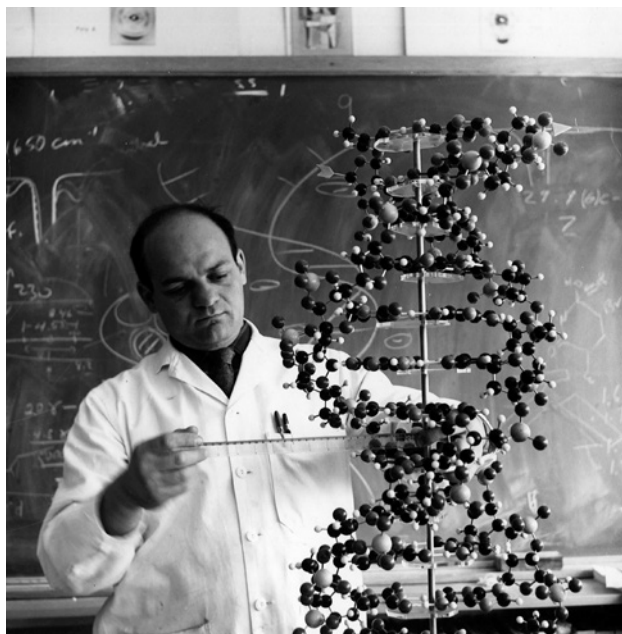
It first came across Alexander Rich in 1963. He was on the cover of that year’s 13 May issue of *Newsweek* with his PhD student Jonathan Warner. The two of them had just discovered clusters of ribosomes called polysomes — crucial components involved in the building of proteins.

Rich, who died on 27 April, was born in 1924 in Hartford, Connecticut to immigrant parents from Russia and Eastern Europe. He grew up during the Great Depression in Springfield, Massachusetts, attending a technical secondary school by day, and working nights at a local rifle factory. At one point, his family went to live at a local YMCA club after being evicted from their home. Against the odds, Rich made it to Harvard University in Cambridge, Massachusetts, graduating with a bachelor’s degree in biochemical sciences in 1947. Two years later, he received a medical degree from Harvard Medical School in Boston.

In 1949, Rich joined chemist Linus Pauling at the California Institute of Technology (Caltech), in Pasadena, where he stayed for five years, and learnt about X-ray crystallography. To his regret, he never published with Pauling; when asked what Rich had achieved during his tenure, Pauling apparently replied: “not much, but he must have learned a lot”.

Indeed he had. From Caltech, Rich went on to lead the physical-chemistry section at the US National Institute of Mental Health (NIMH) in Bethesda, Maryland. In 1955, during a leave period at the Cavendish Laboratory in Cambridge, UK, he and Francis Crick determined the structures of two important proteins: polyglycine II and collagen. Back at the NIMH in 1956, three years after the discovery of DNA’s iconic double helix, Rich and his colleagues discovered that RNA can also form a double helix, and even a three-stranded helical structure. The findings paved the way for studies that showed RNA’s capacity to fold into complex architectures.

In 1958, Rich became an associate professor at the Massachusetts Institute of Technology (MIT) in Cambridge. There he showed that RNA could hybridize with, or bind to, DNA to form a double helix. From the early 1970s, this phenomenon was widely applied



to the identification of DNA sequences, for instance using the northern blot technique. Today, it forms the basis of DNA chips, which are used to measure the expression levels of tens of thousands of genes at once.

Rich’s work on polysomes, carried out in 1963, revealed how active ribosomes — the protein builders of cells — line up along a messenger RNA molecule, like beads on a string. As the ribosomes move along the mRNA, the corresponding amino acids are stitched together to produce proteins. The work established a defining mechanism in protein building.

In 1973, he made the first determination of an RNA double-helix structure at atomic resolution. This was followed, in 1974, by the solution of the L-shaped structure of a transfer RNA molecule, which was made simultaneously by Rich’s MIT group and Aaron Klug’s group at the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK.

Rich is perhaps best known for his discovery of a DNA structure in which the double helix winds to the left instead of the right. In 1979, he, along with crystallographer Andrew Wang, revealed this stable, ‘left-handed’ DNA structure, dubbed Z-DNA, using X-ray crystallography. After showing that Z-DNA can influence the production and alteration of certain mRNA molecules,

Rich worked out the structure of a Z-DNA fragment bound to an RNA-editing enzyme. He and his colleagues also showed how the pathogenicity of the vaccinia virus, and probably of the smallpox virus, correlated with a virus-specific protein binding to the host’s Z-DNA.

Rich’s interest in the latest discoveries across diverse disciplines was irrepressible. During the 1970s, he worked as an adviser for NASA, weighing in on projects exploring the possible existence of life on Mars. He also ventured into biotechnology and co-founded three companies: Repligen, Alkermes, and in his 80s, 3-D Matrix.

Rich received numerous honorary degrees and awards, including the US National Medal of Science, presented to him in 1995 by then US President Bill Clinton.

In spite of such a broad sweep of achievements, Alex was best known among close colleagues for his self-possession, large personality, critical intellect and humanity. He and his wife Jane held legendary parties at their classic brick house near Harvard Square, bringing together all sorts of people, including his four children and now seven grandchildren.

A few years after I saw him on the cover of *Newsweek*, Alex and I became faculty colleagues at MIT. We had endless conversations, ate as often as five times a week at a fish restaurant in Cambridge, and drove around in his dreadful old and enormous cars. On one occasion, the police stopped us, suspecting that his wild gesturing — to explain a theory of evolution to me — indicated drunk driving.

Alex was unstoppable. Once, because of a weeknight family obligation, I had to decline yet another of his dinner invitations. He responded immediately: “No problem, I will come instead to your home later in the evening to talk”. And he did. ■

Paul Schimmel is professor of cell and molecular biology at the Scripps Research Institute in Jupiter, Florida, and La Jolla, California. He was a colleague of Alexander Rich at the Massachusetts Institute of Technology in Cambridge from 1967 onwards.
e-mail: schimmel@scripps.edu

JOSHUA D. RICH

Wild-type microglia do not reverse pathology in mouse models of Rett syndrome

ARISING FROM N. C. Derecki *et al.* *Nature* **484**, 105–109 (2012); doi:10.1038/nature10907

Rett syndrome is a severe neurodevelopmental disorder caused by mutations in the X chromosomal gene *MECP2* (ref. 1), and its treatment so far is symptomatic. *Mecp2* disruption in mice phenocopies

major features of the syndrome² that can be reversed after *Mecp2* re-expression³. Recently, Derecki *et al.*⁴ reported that transplantation of wild-type bone marrow into lethally irradiated *Mecp2*-null

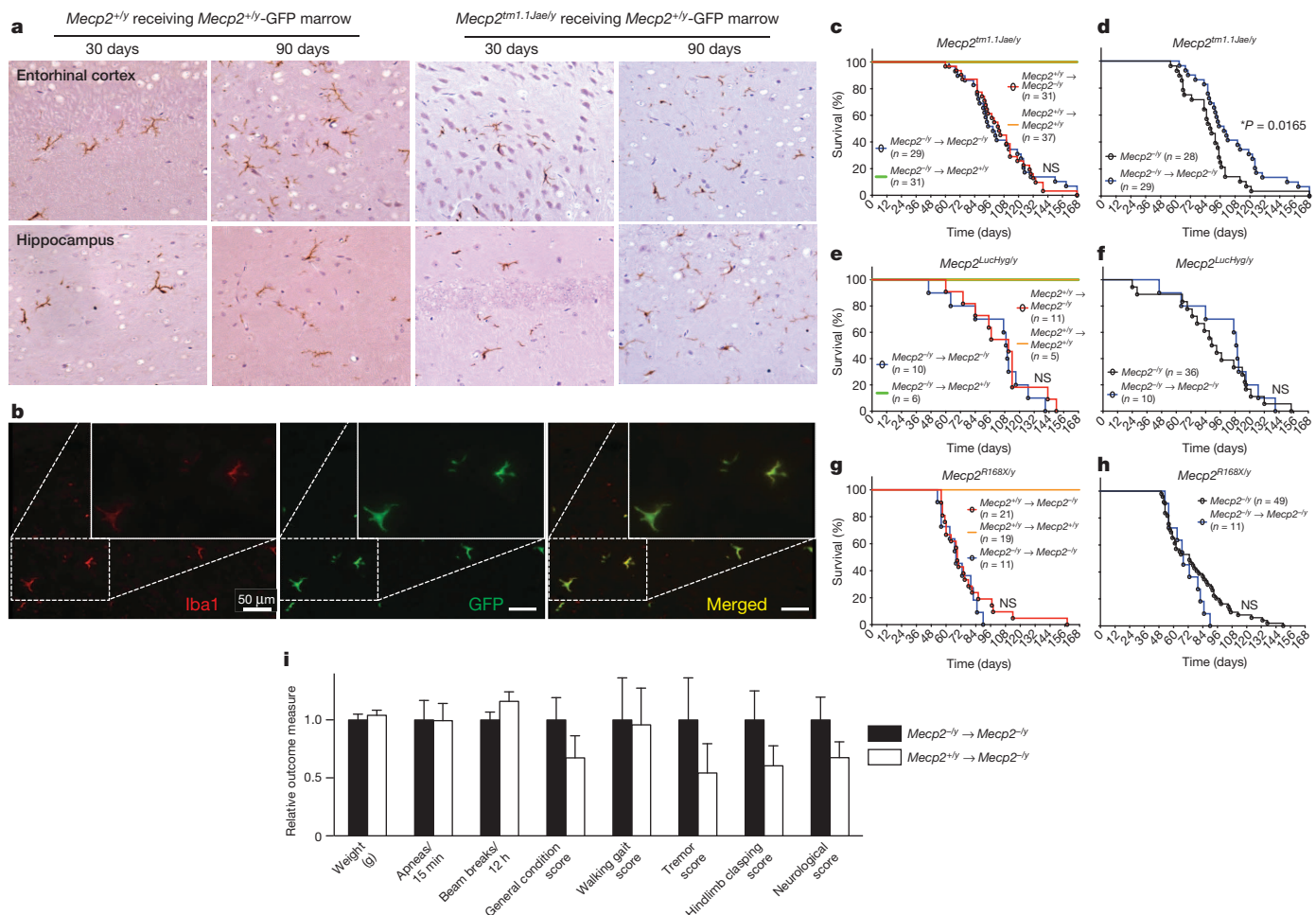


Figure 1 | Early transplantation of wild-type microglia into the brain does not rescue *Mecp2*-null mice. **a**, Transplantation of bone marrow from C57BL/6 (*Mecp2*^{+/y}-GFP marrow) mice with ubiquitous GFP transgene expression into *Mecp2*^{tm1.1Jae/y} mice confers robust donor engraftment at the indicated time after transplant (30 or 90 days) shown via immunohistochemical detection of GFP-positive cells of microglial morphology in entorhinal cortex and hippocampus, in both the *Mecp2*^{+/y}-GFP donors to *Mecp2*^{tm1.1Jae/y} recipient (right panels) and to *Mecp2*^{+/y} recipient (left panels) groups. Original magnification, $\times 400$. **b**, Double immunofluorescent labelling with GFP and the microglia marker Iba1 in cerebellar tissue from *Mecp2*^{+/y}-GFP donors to *Mecp2*^{tm1.1Jae/y} recipient mice examined 30 days after BMT confirms early microglial engraftment in brain parenchyma. Quantification of microglial engraftment is shown in Extended Data Fig. 1c. **c–h**, No differences in survival were noted between transfer of *Mecp2*^{+/y} to *Mecp2*-null (*Mecp2*^{-/-}) mice and of *Mecp2*-null to *Mecp2*-null mice, indicating that engraftment of wild-type microglia into the brains of *Mecp2*-null mice did not protect *Mecp2*-null mice from premature death. In addition, no differences in survival were noted between transfer of *Mecp2*-null to *Mecp2*^{+/y} mice and of

Mecp2^{+/y} to *Mecp2*^{+/y} mice, indicating that engraftment of *Mecp2*-null microglia into the brains of wild-type mice does not shorten survival as seen in *Mecp2*-null mice. NS, not significant. **i**, No differences were seen in other outcome measures at 12 weeks of age (8 weeks after BMT) between *Mecp2*^{+/y} to *Mecp2*^{tm1.1Jae/y} (also termed *Mecp2*^{-/-}) mice (*n* = 31) and *Mecp2*^{tm1.1Jae/y} to *Mecp2*^{tm1.1Jae/y} mice (*n* = 25), including weight, frequency of breathing apneas, locomotor activity (beam breaks), general condition, walking gait, tremor, hindlimb clasping or neurological score. Here, data are presented as relative outcome measure (mean and s.d. for each measure were calculated, and values were divided by the mean value for the *Mecp2*^{tm1.1Jae/y} to *Mecp2*^{tm1.1Jae/y} transplantation mice). Specific values obtained for *Mecp2*^{tm1.1Jae/y} to *Mecp2*^{tm1.1Jae/y} mice and *Mecp2*^{+/y} to *Mecp2*^{tm1.1Jae/y} mice are as follows: weight (in g) (18.22 ± 0.93 versus 18.96 ± 0.8); apneas per 15 min (35.4 ± 5.96 versus 35.2 ± 5.24); beam breaks per 12 h ($3,729.2 \pm 253.3$ versus $4,330.2 \pm 305.2$); general condition score (0.52 ± 0.1 versus 0.35 ± 0.1); walking gait score (0.24 ± 0.087 versus 0.23 ± 0.076); tremor score (0.24 ± 0.087 versus 0.13 ± 0.061); hindlimb clasping score (0.48 ± 0.12 versus 0.29 ± 0.083); neurological score (1.48 ± 0.29 versus 1.0 ± 0.2). None of the differences were statistically significant.

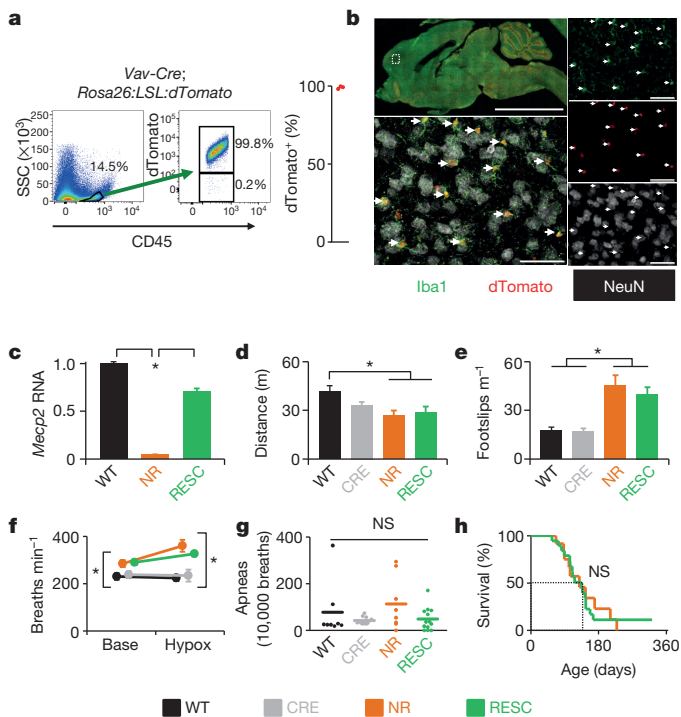


Figure 2 | Genetic reconstitution of *Mecp2* in microglia does not rescue *Mecp2*-null mice. **a, b,** Evaluation of efficiency and specificity of *Vav1-Cre* in microglia. **a,** Representative flow sorting of microglia derived from brains from *Vav1-Cre; Rosa26:LSL:tdTomato* mice shown on left, with quantification ($n = 3$) of cells that express both fluorescent reporter (tdTomato) and microglia marker (CD45) shown on right. SSC, side scatter. **b,** Histological characterization of tdTomato expression in brain from *Vav1-Cre; Rosa26:LSL:tdTomato*. Top left, low-power image of a mid-sagittal section; bottom left, higher power image of cortex. Right, individual colour channels contributing to merged image (bottom left). Arrows indicate microglia expressing tdTomato (Iba1⁺/tdTomato⁺). Scale bars, 5 mm (low power) and 50 μ m (high power). **c,** Quantitative PCR (qPCR) results of *Mecp2* expression from flow-sorted microglia derived from wild-type (WT; $n = 2$), *Mecp2*^{LSL/Y} (NR; $n = 2$) and *Vav1-Cre*^{Tg/+}; *Mecp2*^{LSL/Y} (RESC; $n = 3$) animals. **d,** Distance travelled in open field assay. CRE denotes *Vav1-Cre*^{Tg/+} mice. **e,** Number of footslips per distance travelled on parallel rods. **f,** Breathing rate at baseline and during hypoxia challenge. **g,** Number of apneas per 10,000 breaths. **h,** Survival curve. In **d–h**, WT $n = 8$; CRE $n = 10$; NR $n = 12$; RESC $n = 13$. Data are mean and s.e.m. * $P < 0.05$. Statistical analyses in **c–g** analysed by one-way analysis of variance (ANOVA) with post-hoc pairwise *t*-test with Bonferroni correction, and in **h**, Kaplan–Meier survival analysis was used.

(*Mecp2*^{tm1.1Jae/y}) mice prevented neurological decline and early death by restoring microglial phagocytic activity against apoptotic targets⁴, and clinical trials of bone marrow transplantation (BMT) for patients with Rett syndrome have thus been initiated⁵. We aimed to replicate and extend the BMT experiments in three different Rett syndrome mouse models, but found that despite robust microglial engraftment, BMT from wild-type donors did not prevent early death or ameliorate neurological deficits. Furthermore, early and specific *Mecp2* genetic expression in microglia did not rescue *Mecp2*-deficient mice.

We first sought to replicate BMT-mediated rescue of male mice derived from the same *Mecp2*^{tm1.1Jae/y} colony used in the original report⁴, implementing established standards for conducting preclinical studies^{2,6}. Mice were maintained on a C57Bl/6J background, which was confirmed in recipient animals by genome scanning (see Supplementary Information). Four-week-old *Mecp2*^{tm1.1Jae/y} mice and wild-type littermates were subjected to the same protocol of lethal

split-dose γ -irradiation. Mice were then randomized to receive tail vein injection of bone marrow from *Mecp2*-deficient male littermates or from *Mecp2*-proficient animals, including C57Bl/6J male mice ubiquitously expressing green fluorescent protein (GFP) and *Mecp2*^{+/y} littermates of the recipients. All animals achieved multilineage peripheral blood engraftment as judged by the fraction of donor-derived GFP-expressing cells in peripheral blood 4 and 8 weeks after transplant (Extended Data Fig. 1a). PCR analysis of blood and tail tissue 4 weeks after transplant also confirmed expression of the appropriate mutant or wild-type variant of *Mecp2* in blood in all groups (Extended Data Fig. 1b). Microglial engraftment in brain parenchyma 30 and 90 days after transplant was similar in mutant and wild-type recipients engrafted with marrow from wild-type mice ubiquitously expressing a GFP transgene (Fig. 1a, b and Extended Data Fig. 1c), and comparable to engraftment observed by Derecki *et al.*⁴ and others⁷.

Contrary to our expectation, *Mecp2*^{tm1.1Jae/y} mice that received *Mecp2*^{+/y} marrow had no extension of lifespan compared to *Mecp2*^{tm1.1Jae/y} marrow recipients (Fig. 1c). No difference in survival was observed in mutant animals that received *Mecp2*^{+/y} marrow from wild-type littermates or C57Bl/6J animals ubiquitously expressing GFP (Extended Data Fig. 1d). We also observed no benefit in outcome measures at 12 weeks of age, 8 weeks after transplant, including weight, breathing, locomotion, general condition, walking gait, tremor, hindlimb clasping or neurological score (Fig. 1i). Thus, the same BMT procedure with substantially greater numbers of animals, randomly assigned to treatment group, with mice from the same *Mecp2*^{tm1.1Jae/y} colony did not replicate any aspects of the protective effect reported by Derecki *et al.*⁴. Furthermore, histological analysis blind to genotype and treatment group showed no neuropathological evidence of differential apoptosis, microglial response, or tissue degeneration between experimental groups (Extended Data Fig. 1e). There was also no protective effect on survival after BMT in two additional mouse models of Rett syndrome (Fig. 1e, g): *Mecp2*^{LucHyg/y} mice that contain a *Mecp2*-firefly luciferase/hygromycin-resistance gene fusion (Extended Data Fig. 2a–e) and *Mecp2*^{R168X/y} mice⁸, despite excellent engraftment after BMT (Extended Data Fig. 2f–h). Experiments with these two models were performed in independent laboratories following the same BMT protocol⁴.

In all models, wild-type mice transplanted with wild-type bone marrow showed no mortality, indicating that the procedure was well tolerated (Fig. 1c, e, g). Likewise, BMT was well-tolerated by mutant animals, as *Mecp2* mutant animals receiving mutant marrow exhibited either no change (*Mecp2*^{LucHyg/y} and *Mecp2*^{R168X/y} mice), or, surprisingly, slightly reduced mortality (*Mecp2*^{tm1.1Jae/y} mice) compared to naive mice not subjected to BMT (Fig. 1d, f, h). The small survival extension may be related to a salutary effect of post-irradiation antibiotic treatment of transplanted animals, to which naive animals were not exposed, or to differences in animal handling⁹.

To address the role for microglia in Rett syndrome reported by Derecki *et al.*⁴ further, we used the *Cre/lox* system and a *lox-stop-lox* allele of *Mecp2* (*Mecp2*^{LSL}, referred to as *Mecp2*^{lox-stop/y} in ref. 4) to examine the effect of genetically driven expression of *Mecp2* in microglia during development (see Supplementary Information for full Methods details). First, we analysed the suitability of the *LysM-Cre* transgene (*LysM*^{Cre} in ref. 4; *LysM* is also known as *Lyz2*), which was used by Derecki *et al.*⁴ in their genetic *Mecp2*^{LSL/y} rescue experiments⁴, to drive efficient microglia-specific gene restoration. As previously reported¹⁰, *LysM-Cre*-driven dTomato reporter cells accounted for less than 25% of microglia, as assessed using flow cytometry of microglia derived from mice containing the *LysM-Cre* transgene and a transgene expressing Cre-dependent dTomato (Extended Data Fig. 3a). Furthermore, when

we generated *LysM-Cre; Mecp2^{LSL/Y}* mice (termed *Mecp2^{lox-stop/y} LysM^{cre}* in ref. 4), we observed MeCP2 expression in neurons (large NeuN⁺ cells) in many brain regions (Extended Data Fig. 3b).

To identify a Cre transgenic line that drives efficient expression within microglia, we next evaluated the *Vav1-Cre* transgene, which selectively expresses throughout the haematopoietic compartment¹¹. In contrast to *LysM-Cre*, the *Vav1-Cre* transgene targeted microglia with high efficiency (Fig. 2a) and specificity (Fig. 2b). As *Vav1-Cre*-driven expression in brain proved to be efficient and restricted to microglia, we applied this system to test whether expression of *Mecp2* in microglia rescues *Mecp2*-null mice. To quantify *Mecp2* restoration in microglia, we used the *fms-GFP* transgene, the expression of which within brain is restricted to microglia, for flow sorting¹¹ (Extended Data Fig. 3c). Microglia derived from *Vav1-Cre; Mecp2^{LSL/Y}* animals expressed *Mecp2* messenger RNA at 75% of the level of *Mecp2* mRNA in microglia derived from *Mecp2^{+/Y}* animals (Fig. 2c). Similar to other *Mecp2*-null mouse models, *Mecp2^{LSL/Y}* animals showed hypoactivity, poor motor coordination on parallel rod walking, increased basal and hypoxia breathing rate, increased frequency of apneas, and early death, none of which was improved by *Mecp2* expression in microglia of *Vav1-Cre; Mecp2^{LSL/Y}* animals (Fig. 2d–h). We thus conclude that driving *Mecp2* expression developmentally in microglia did not ameliorate the phenotype of *Mecp2*-null mice, in contrast to the data reported by Derecki *et al.*⁴.

In conclusion, our experiments do not support BMT as therapy for Rett syndrome. We observe no benefit of BMT-mediated delivery of wild-type microglia into the brains of three different preclinical models of Rett syndrome, nor do we observe a causative role of microglia in the disease process. Our BMT studies included large numbers of mice derived from the same parent colony used in the original report⁴, with treatment assigned randomly and analysis conducted blind to genotype and treatment group. Finally, we showed that even early and highly efficient genetically driven *Mecp2* expression in the microglia of *Mecp2*-null mice conferred no protective effect. Restoration of MECP2 in microglia through either BMT or genetics did not rescue the major observed phenotypes in Rett syndrome, which argues against the previously proposed therapeutic potential of BMT in patients with Rett syndrome⁴.

Jieqi Wang^{1*}, Jan Eike Wegener^{2*}, Teng-Wei Huang^{3,4*}, Smitha Sripathy⁵, Hector De Jesus-Cortes^{6,7}, Pin Xu⁶, Stephanie Tran¹, Whitney Knobbe¹, Vid Leko⁵, Jeremiah Britt⁷, Ruth Starwalt⁸, Latisha McDaniel⁷, Chris S. Ward³, Diana Parra³, Benjamin Newcomb⁵, Uyen Lao⁵, Cynthia Nourigat⁵, David A. Flowers⁵, Sean Cullen⁴, Nikolas L. Jorstad⁹, Yue Yang⁹, Lena Glaskova⁹, Sebastian Vigneau¹⁰, Julia Kozlitina¹¹, Michael J. Yetman¹², Joanna L. Jankowsky¹², Sybille D. Reichardt¹³, Holger M. Reichardt¹³, Jutta Gärtner², Marisa S. Bartolomei¹⁰, Min Fang^{5,9}, Keith Loeb^{5,9}, C. Dirk Keene⁹, Irwin Bernstein⁵, Margaret Goodell^{4,14,15,16,17,18}, Daniel J. Brat¹⁹, Peter Huppke²⁵, Jeffrey L. Neul^{3,4,12,14,16,17,18,§}, Antonio Bedalov^{5,20,§} & Andrew A. Pieper^{7,21,22,23,§}

¹Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

²Department of Pediatrics and Adolescent Medicine, Division of Pediatric Neurology, University Medical Center Göttingen, Robert-Koch-Strasse 40, 37075 Göttingen, Germany.

email: phuppke@med.uni-goettingen.de

³Jan and Dan Duncan Neurological Research Institute (Texas Children's Hospital), Baylor College of Medicine, Houston, Texas 77030, USA.

email: jneul@ucsd.edu

⁴Program in Developmental Biology, Baylor College of Medicine, Houston, Texas 77030, USA.

⁵Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA.

email: abedalov@fredhutch.org

⁶Graduate Program of Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

⁷Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA.

email: andrew-pieper@uiowa.edu

⁸Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

⁹Department of Pathology, University of Washington School of Medicine, Seattle, Washington 98195, USA.

¹⁰Department of Cell & Developmental Biology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

¹¹Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

¹²Department of Neuroscience, Baylor College of Medicine, Houston, Texas 77030, USA.

¹³Institute for Cellular and Molecular Immunology; University of Göttingen Medical School, Humboldtallee 34, 37073 Göttingen, Germany.

¹⁴Department of Pediatrics, Baylor College of Medicine, Houston, Texas 77030, USA.

¹⁵Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, Texas 77030, USA.

¹⁶Stem Cell and Regenerative Medicine Center, Baylor College of Medicine, Houston, Texas 77030, USA.

¹⁷Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA.

¹⁸Program in Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, Texas 77030, USA.

¹⁹Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia 30322, USA.

²⁰Department of Medicine, University of Washington School of Medicine, Seattle, Washington 98105, USA.

²¹Department of Neurology, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA.

²²Veterans Affairs, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA.

²³Weill Cornell Autism Research Program, Weill Cornell Medical College, New York, New York 10065, USA.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

Received 9 September 2014; accepted 11 March 2015.

- Neul, J. L. The relationship of Rett syndrome and MECP2 disorders to autism. *Dialogues Clin. Neurosci.* **14**, 253–262 (2012).
- Katz, D. M. *et al.* Preclinical research in Rett syndrome: setting the foundation for translational success. *Dis. Model. Mech.* **5**, 733–745 (2012).
- Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315**, 1143–1147 (2007).
- Derecki, N. C. *et al.* Wild-type microglia arrest pathology in a mouse model of Rett syndrome. *Nature* **484**, 105–109 (2012).
- MT2013-31: Allogeneic hematopoietic cell transplantation for inherited metabolic disorders, severe osteoporosis and males with Rett syndrome following conditioning with busulfan (therapeutic drug monitoring), fludarabine +/- ATG; Masconic Cancer Center, University of Minnesota. ClinicalTrials.gov identifier: NCT02171104 (2014).
- Landis, S. C. *et al.* A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* **490**, 187–191 (2012).
- Yang, Y. *et al.* Perivascular, but not parenchymal, cerebral engraftment of donor cells after non-myeloablative bone marrow transplantation. *Exp. Mol. Pathol.* **95**, 7–17 (2013).
- Brendel, C. *et al.* Readthrough of nonsense mutations in Rett syndrome: evaluation of novel aminoglycosides and generation of a new mouse model. *J. Mol. Med. (Berl.)* **89**, 389–398 (2011).

9. Pitcher, M. R. *et al.* Insulinotropic treatments exacerbate metabolic syndrome in mice lacking MeCP2 function. *Hum. Mol. Genet.* **22**, 2626–2633 (2013).
10. Goldmann, T. *et al.* A new type of microglia gene targeting shows TAK1 to be pivotal in CNS autoimmune inflammation. *Nature Neurosci.* **16**, 1618–1626 (2013).
11. Sasmono, R. T. *et al.* A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse. *Blood* **101**, 1155–1163 (2003).
12. Chen, R. T., Akbarian, S., Tudor, M. & Jaenisch, R. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nature Genet.* **27**, 327–331 (2001).

Supplementary Information is available in the online version of the paper.

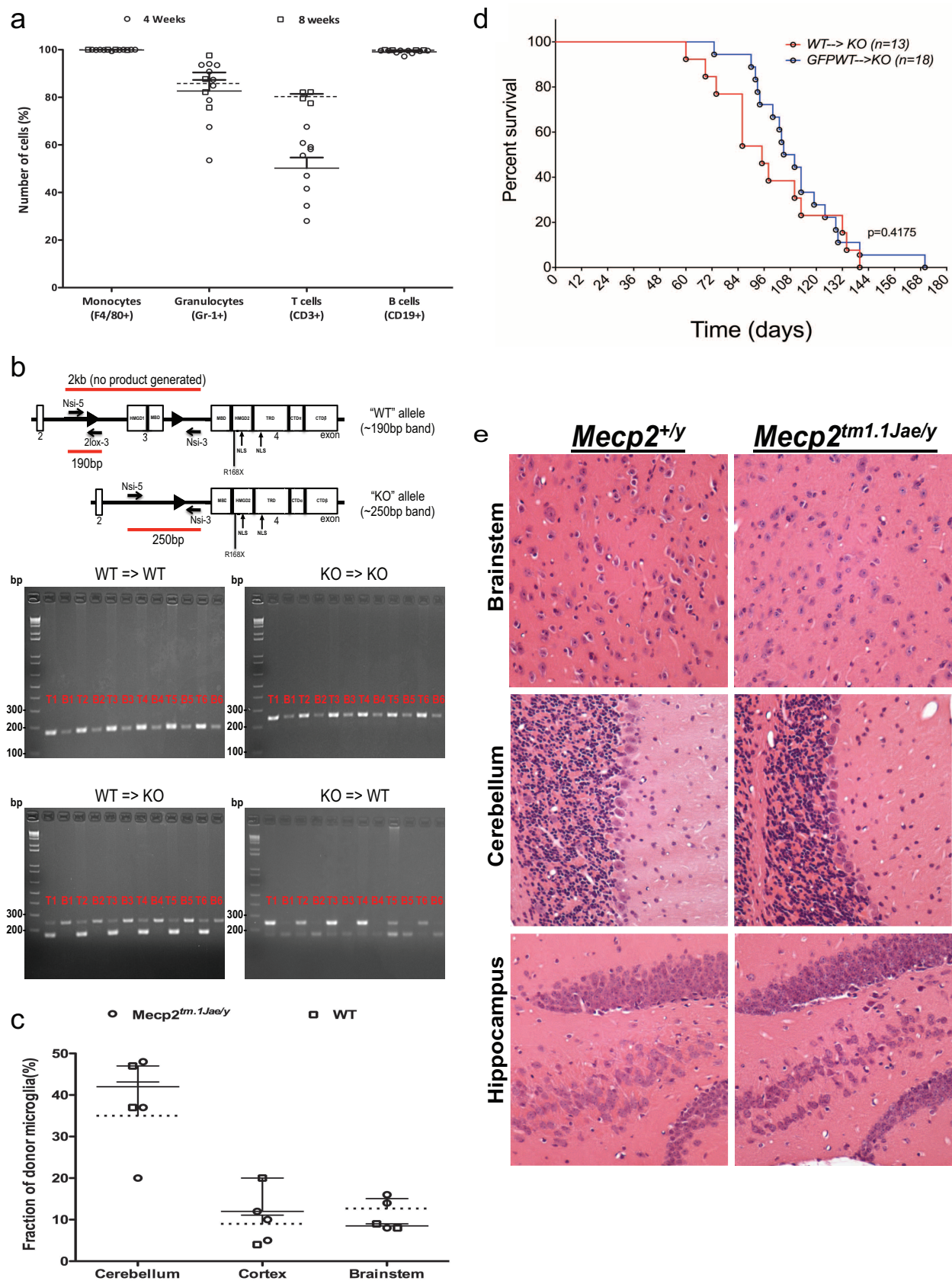
Author Contributions All experiments were conceived of and supervised by P.H., J.L.N., A.B. and A.A.P., who are jointly responsible for the work presented

and are corresponding authors. Designed experiments: A.A.P., A.B., S.S., V.L., J.W., P.X., R.S., D.J.B., J.L.N., T.-W.H., C.S.W., D.P., S.C., M.G., J.E.W., P.H., M.J.Y., J.L.J., S.D.R., H.M.R., J.G. Collected data: B.N., C.N., U.L., D.A.F., N.L.J., Y.Y., S.V., L.G., J.W., P.X., S.T., W.K., J.B., R.S., L.M., D.J.B., T.-W.H., C.S.W., D.P., S.C., J.E.W., M.J.Y., J.L.J., S.D.R., H.M.R. Analysed data: A.A.P., A.B., J.W., K.L., M.F., C.D.K., I.B., H.D.J.-C., P.X., R.S., D.J.B., J.L.N., T.-W.H., C.S.W., D.P., S.C., M.G., J.E.W., P.H., S.D.R., H.M.R., J.G. Wrote manuscript: A.A.P., A.B., J.W., D.J.B., H.D.J.-C., J.L.N., T.-W.H., P.H., J.E.W. Edited manuscript: A.A.P., A.B., M.S.B., C.D.K., I.B., H.D.J.-C., D.J.B., J.L.N., T.-W.H., C.S.W., D.P., S.C., M.G., P.H., J.G., H.M.R.

Competing Financial Interests Declared none.

doi:10.1038/nature14444

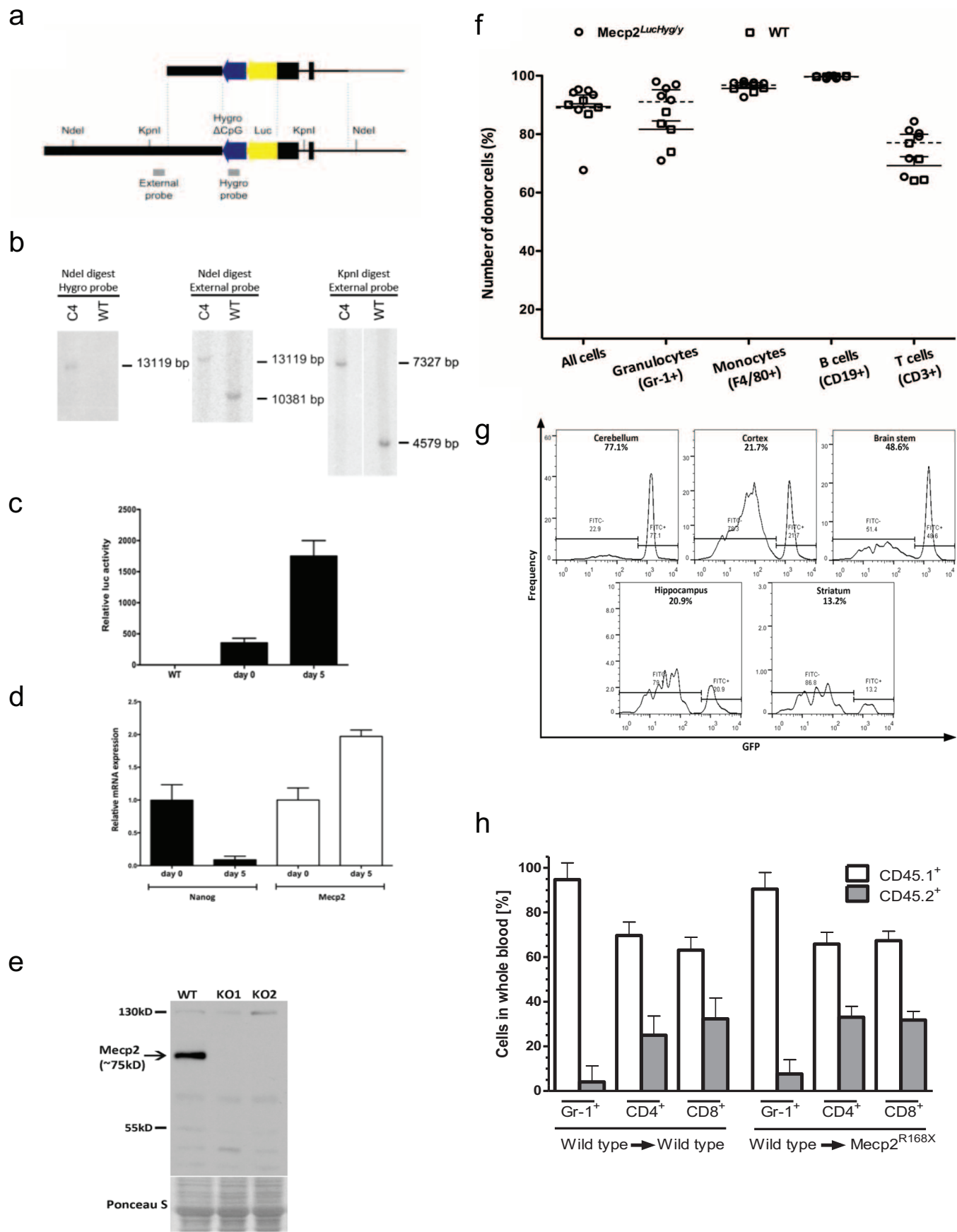
BRIEF COMMUNICATIONS ARISING



Extended Data Figure 1 | Engraftment with donor cells after bone marrow transplantation, and lack of evidence of neuropathology in *Mecp2*-null animals. **a**, Multilineage peripheral blood engraftment with donor cells in *Mecp2*^{tm1.1Jae/y} and wild-type mice. Wild-type and *Mecp2*^{tm1.1Jae/y} animals received transplant from wild-type animals ubiquitously expressing GFP (Jackson Labs, C57BL/6-Tg(UBC-GFP)20Scha/J, stock 004353). Peripheral blood engraftment in indicated blood lineages was measured by flow cytometry (GFP) 4 and 8 weeks after transplant. **b**, PCR analysis of blood and tail tissue 4 weeks after transplant. Expression of only the appropriate mutant or wild-type variant of *Mecp2* from the donor in blood in all four groups is shown, with retention of the original genotype in tail tissue as expected. Specifically, *Mecp2*^{+/y} to *Mecp2*^{+/y} mice show only the wild-type allele at 190 base pairs (bp), whereas *Mecp2*^{tm1.1Jae/y} to *Mecp2*^{tm1.1Jae/y} mice show only the mutant allele at 250 bp, as previously described in the original report of generation of these mice¹². *Mecp2*^{tm1.1Jae/y} to *Mecp2*^{+/y} mice, however, show only the mutant allele in blood tissue and retention of host wild-type allele in tail tissue. Accordingly, *Mecp2*^{+/y} to *Mecp2*^{tm1.1Jae/y} mice show only the wild-type allele in blood tissue, with retention of the host mutant allele in tail tissue. Tail tissue in these latter two groups shows some of the allele from the donor as well, presumably owing to blood contained within the tail clips used for analysis. Notably, the *Mecp2* allele expressed in blood is always restricted to the donor genotype, indicating successful transplantation with complete replacement of the haematopoietic system in the host. Samples are labelled with a 'T' for tail and 'B' for blood, followed by the number of the animal, indicating that six different animals were analysed for each condition. CTD, C terminus domain α

and β ; HMGD1/2, high mobility group protein-like domain 1/2; MBD, methyl binding domain; NLS, nuclear localization signal; TRD, transcription repression domain. **c**, Robust and early microglial engraftment of donor cells after BMT in *Mecp2*^{tm1.1Jae/y} and wild-type mice. Microglial engraftment was visualized using double immunofluorescence staining in sections quenched for autofluorescence by incubation in Sudan black solution. All sections were stained with an anti-Iba1 primary with CY-3 secondary and an anti-GFP primary with CY-5 secondary. All microglia are Iba1-positive, and thus successfully engrafted GFP-expressing donor-derived microglia were observed as GFP⁺/Iba1⁺, whereas native microglia were only Iba1⁺. Engraftment of microglia into wild-type and *Mecp2*^{LuchTyg/y} mice was determined by dividing the GFP⁺/Iba1⁺ cells by the number of total Iba1⁺ cells. Cell counts were performed in cerebellum, cortex and brainstem from mice. Percentage engraftment in wild-type and *Mecp2*^{tm1.1Jae/y} mice yielded similar results to previously published engraftment results at 30 days after transplantation⁹. **d**, BMT was well-tolerated in animals. No difference in survival was observed in mutant animals that received *Mecp2*^{+/y} marrow from their wild-type littermates ($n = 13$) and C57Bl/6J animals ubiquitously expressing GFP ($n = 13$). KO, knockout. **e**, Representative haematoxylin-and-eosin-stained sections of cerebellum, brainstem and hippocampus from age-matched wild-type and *Mecp2*^{tm1.1Jae/y} mice killed at 7 weeks of age. Original magnification, $\times 400$. Sections demonstrate comparable histological features between wild-type and *Mecp2*^{tm1.1Jae/y} brains, and a lack of gliosis, cell loss, cellular debris, microglia or macrophages in *Mecp2*^{tm1.1Jae/y} brains.

BRIEF COMMUNICATIONS ARISING

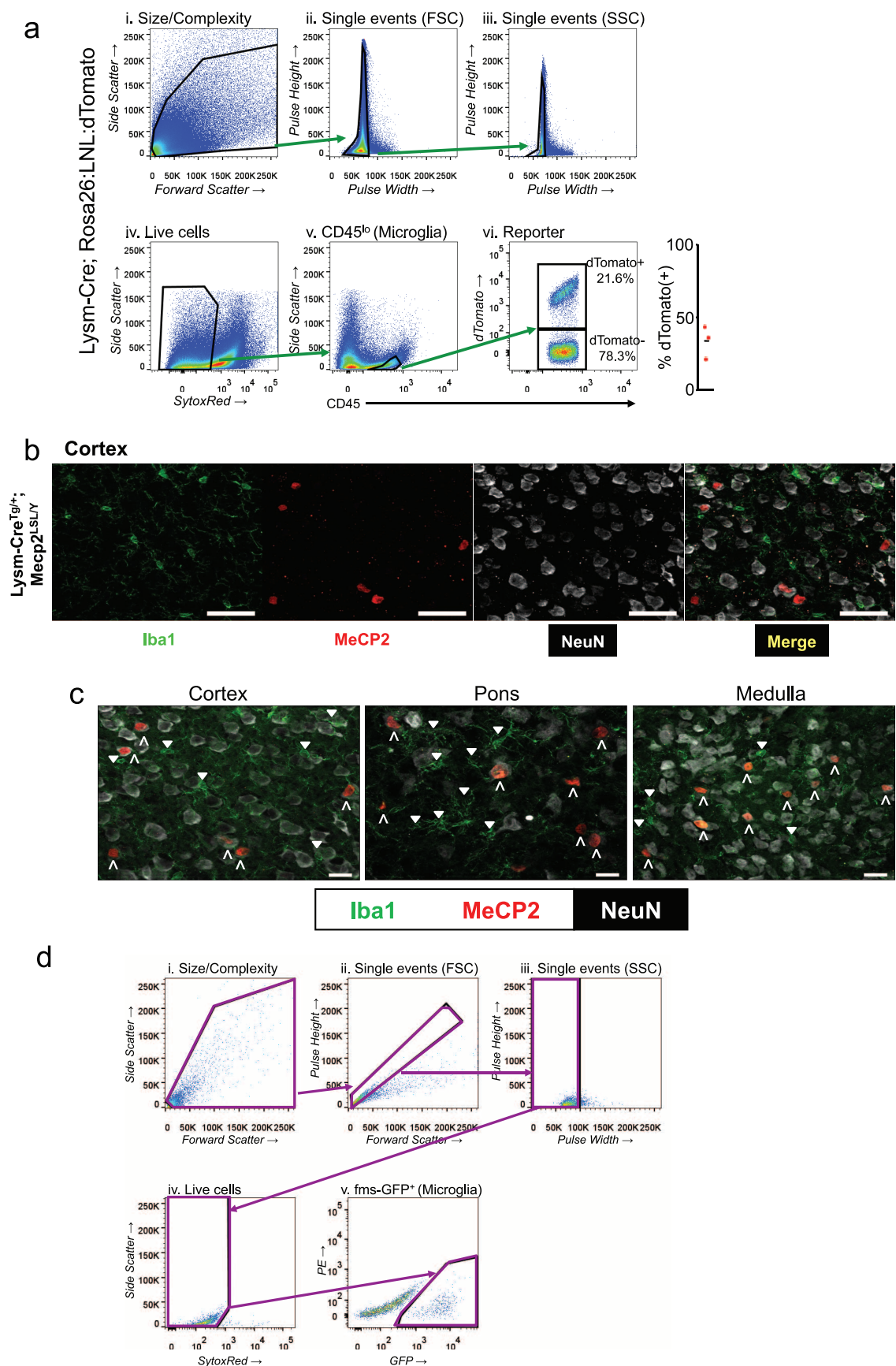


BRIEF COMMUNICATIONS ARISING

Extended Data Figure 2 | Early transplantation of wild-type microglia into the brain does not rescue additional models of *Mecp2*-null mice:

***Mecp2*^{LucHyg} mice and C57Bl/6J *Mecp2*^{R168X} mice.** **a**, Generation of *Mecp2*^{LucHyg} mice. Luciferase/hygromycin (LucHyg) fusion gene vector correctly targeted to the *Mecp2* locus in embryonic stem cells. Positions of the probes and enzyme restriction sites are indicated. The homology arms of the targeting vector are depicted in black, and its backbone in grey. **b**, Confirmation of genetic targeting for *Mecp2*^{LucHyg} mice. Southern blotting of NdeI- or KpnI-digested DNA extracted from clone C4 cells, used for blastocyst injections, hybridized with either the hygromycin or external probe confirms correctly targeted event. **c**, Luciferase activity in clone C4 cells before (day 0) or after (day 5) subjecting cells to retinoic-acid-induced differentiation. After adsorption to eliminate feeder mouse embryonic fibroblasts, clone C4 embryonic stem cells were treated with retinoic acid (100 nM) in differentiation medium for 5 days, and luciferase activity was measured before and after retinoic acid treatment. Mean values are plotted relative to that of the wild-type cells ($n = 3$, error bars denote s.d.). Retinoic-acid-induced differentiation leads to an increase in luciferase activity consistent with an increase in *Mecp2* expression level as measured in **d**. **d**, mRNA levels of *Mecp2* increased and of embryonic stem-cell marker *Nanog* decreased in clone C4 cells subjected to retinoic-acid-induced differentiation. mRNA levels were

measured before and after treatment by qPCR. Mean values plotted relative to day 0 for each mRNA ($n = 3$, error bars denote s.d.). **e**, Western blot analysis of MECP2 expression in brains of wild-type and *Mecp2*^{LucHyg} male mice. MECP2 protein is not detected in MECP2 luciferase males. Ponceau S staining serves as a loading control. **f, g**, Robust peripheral blood and microglial engraftment of donor cells after BMT in *Mecp2*^{LucHyg/y} mice. Wild-type and *Mecp2*^{LucHyg/y} mice received wild-type bone marrow marked with GFP or CD41.1. Peripheral blood engraftment was measured by flow cytometry (GFP or CD41.1) in the indicated lineage 4–8 weeks after transplantation. For central nervous system engraftment, flow cytometry was performed on isolated mononuclear cells from the cortex, brainstem, cerebellum, hippocampus and striatum. Engraftment of BMT-derived cells was determined by dividing the CD11b⁺CD45⁺GFP⁺ cell population by total CD11b⁺CD45⁺ monocytes/microglia. **h**, Robust peripheral blood engraftment of donor cells 7 weeks after BMT in *Mecp2*^{R168X} mice. Reconstitution of bone marrow from *B6.SJL-Ptprca*^a *Pepe*^b/*BoyJ* mice into wild-type mice and C57Bl/6J *Mecp2*^{R168X} mice showed robust engraftment in peripheral blood. Reconstitution of bone marrow was determined by FACS analysis of peripheral blood using anti-GR-1, anti-CD4 and anti-CD8 antibodies and CD45.1 for the donor cells (*B6.SJL-Ptprca*^a *Pepe*^b/*BoyJ* mice, white bars) and CD45.2 for host cells (wild-type and C57Bl/6J *Mecp2*^{R168X} mice, grey bars).



BRIEF COMMUNICATIONS ARISING

Extended Data Figure 3 | Flow sorting and histological characterization of *LysM-Cre* or *Vav1-Cre* transgenic mice. **a**, Stepwise process to characterize the amount of microglia (CD45 lo expressing) cells that also express tdTomato in a *LysM-Cre*-dependent fashion. **b**, High power images of cortex from *LysM-Cre*^{Tg/+}; *Mecp2*^{L^{SL}/Y} animals. Scale bars, 50 μ m. **c**, Merged high power images from cortex, pons and medulla from *LysM-Cre*; *Mecp2*^{L^{SL}/Y} animal. Circumflex (^) symbols identify large NeuN staining cells that express MECP2 (NeuN⁺/MECP2⁺); downward-facing triangles mark microglia not expressing MECP2

(Iba1⁺/MECP2⁻). Scale bars, 20 μ m. **d**, Gating strategy for microglia sorting for *Mecp2* expression quantification in *Vav1-Cre*; *Mecp2*^{L^{SL}/Y} and control animals is presented: (i) size/complexity (size/cytoplasmic granularity for cells but not debris); (ii) forward scatter pulse height/area (eliminates doublet cells); (iii) side scatter pulse height/width (eliminates doublet cells); (iv) SYTOX red staining; dead cells are SYTOX-red-positive and removed from the following analysis; (v) *fms-GFP* expression analysis enables the purification of microglia.

Doubtful pathways to cold tolerance in plants

ARISING FROM A. E. Zanne *et al.* *Nature* **506**, 89–92 (2014); doi:10.1038/nature12872

Zanne *et al.*^{1,2} addressed an important evolutionary question: how did flowering plants repeatedly enter cold climates? Herbaceous growth, deciduous leaves, and narrow water-conducting cells are adaptations to freezing. Using phylogenetic analyses, they concluded that herbs and narrow conduits evolved first in the tropics (“trait first”), facilitating movement into freezing areas, but that deciduous leaves evolved in response to freezing temperatures (“climate first”). Unfortunately, after correcting for an error that we uncovered³, the “striking findings” of Zanne *et al.*¹ seem inconclusive; here we highlight methodological issues of more general interest and question the value of their approach. There is a Reply to this Brief Communication Arising by Zanne, A. E. *et al.* *Nature* **521**, <http://dx.doi.org/10.1038/nature14394> (2015).

Zanne *et al.*¹ chose methods that required transforming quantitative variables into binary characters; not surprisingly, we found that their results are highly sensitive to how characters are scored. This is not inherently problematic, but the delineations must be well justified. While we have concerns with each of their thresholds, the climate character underlying their analyses merits special scrutiny. Of the species Zanne *et al.*¹ studied, 50% were represented by collections from both freezing and non-freezing areas; these were scored as “freezing-exposed” if 2.5% of the collections experienced a minimum

temperature of 0 °C. This cut-off used the tail end of species distributions to delineate character states, where we expect considerable error in the data^{4,5}, especially for the many poorly collected species in their sample. Using more stringent data cleansing and/or alternative thresholds for “freezing-exposed”, we obtained a wide range of results (Fig. 1a). For instance, when we required half of the collection sites to experience freezing, the leaf phenology result shifted from 36.7% to 72.5% trait first. Depending on how climate data were handled, results for plant habit varied from 25.3% to 95.5% trait first (see https://github.com/ejedwards/reanalysis_zanne2014) and, contrary to Zanne *et al.*¹, we sometimes found that growth form was twice as evolutionarily labile as climate occupancy.

But our concerns run deeper. Their evolutionary trajectories were inferred using a newly developed method whose behaviour is unexplored. In simulations we discovered that their method strongly infers a preferred trajectory even when none is present (see https://github.com/ejedwards/reanalysis_zanne2014). When the simulated data contained an equal number of climate-first and trait-first transitions, their method inferred a strong climate-first or trait-first trajectory 77% of the time (Fig. 1b). Thus, the preferred trajectories of Zanne *et al.*¹ could have nothing to do with what actually happened during angiosperm evolution, and no attempt was made to connect their

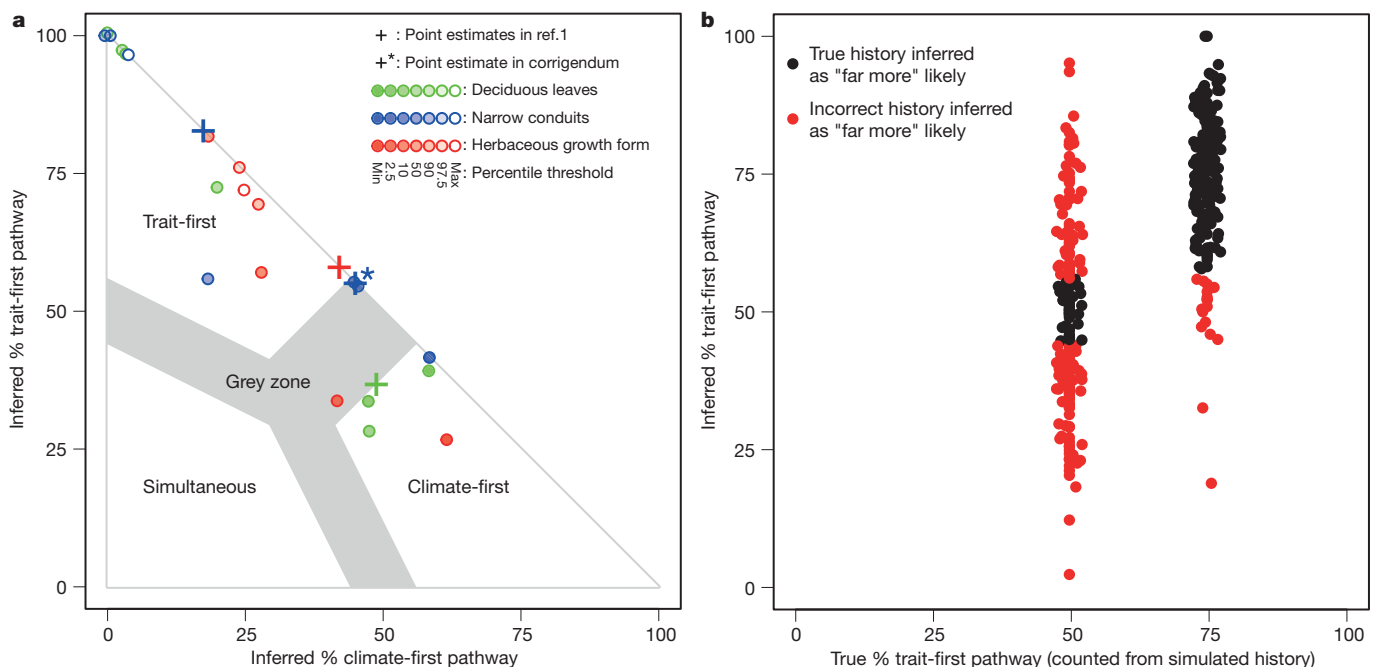


Figure 1 | Unreported uncertainty and potential error. Based on the analyses by Zanne *et al.*¹. **a**, Sensitivity of the Zanne *et al.*¹ results to alternative treatments of the climate data. For each of the three traits, the Zanne *et al.*¹ result is marked by a cross (48.7% “climate first” for deciduousness; 82.7% and 58.0% “trait first” for conduit size and growth form, respectively). Our re-analysis of conduit size using the correct diameter variable shifts their inference to 53.5% trait first (marked with an asterisk), moving their only strong result into a region of questionable significance (the “grey zone”), along with their other two pathways. For each trait we obtained a wide range of outcomes, including apparently decisive support for climate first or trait first, by simple modifications of the climate variable (see Methods). This figure includes only our implementation of different freezing thresholds; for the effect of alternative

data cleansing see https://github.com/ejedwards/reanalysis_zanne2014.

b, Error rates using the Zanne *et al.*¹ transition-rates method. We simulated character evolution with a strongly biased pathway (3 times more “trait-first” transitions) and with no preferred pathway (equal number of “trait-first” and “climate-first” transitions) to examine the behaviour of their method. When there was a strong underlying trajectory in the data, their method could usually detect it. However, when there was no dominant trajectory, their method performed poorly, incorrectly inferring a strongly preferred pathway 77% of the time. Zanne *et al.*¹ described deciduous leaves as being “far more likely” to have evolved climate first (49% vs 37%); on this basis we considered one pathway “far more likely” than another if the difference was 12% or more.

trajectories to inferred character changes in the phylogeny. Consequently, it is unclear, even roughly, how many tropical-to-temperate transitions were sampled, or how their trajectories relate to the directionality of change. In the case of their erroneously binned conduits, for example, the few taxa they scored as having the supposed ancestral condition were instead recently derived within the tree.

Finally, we disagree with Zanne *et al.*'s claim that their results are "qualitatively the same" after correcting for their error. 54% and 83% seem like very different answers, and all of their preferred pathways now hover around a grey zone where their probability is hardly greater than the alternatives (Fig. 1a). In the end, we struggle even to understand the meaning of a number like 54%. It should not be taken to mean that 54% of transitions were trait first when, as we have demonstrated, their method cannot accurately infer the true evolutionary history. Nor should we interpret their result as if every species had a 54% chance of a trait-first transition, when their own sub-analyses of growth form showed that these probabilities vary widely by clade. We urge greater caution in conducting and interpreting phylogenetic analyses at this scale, and believe that it will be more productive to focus instead on concrete, carefully developed case studies that incorporate more of the relevant variables⁶.

Methods

We employed seven different thresholds to define a species as "freezing exposed", using various percentiles of localities experiencing 0 °C. We excluded duplicate records, enforced minimum sample sizes ($n = 3$, $n = 10$), performed alternative data grooming procedures, and re-ran the original analyses across all data sets. We also simulated character histories with differing degrees of bias towards particular pathways. We scored the relative frequency of the trait-first pathway

from each simulation, and compared it to the trait-first probability inferred using their method. Annotated scripts and analyses are publicly archived in https://github.com/ejedwards/reanalysis_zanne2014.

Erika J. Edwards¹, Jurriaan M. de Vos¹ & Michael J. Donoghue²

¹Department of Ecology and Evolutionary Biology, Brown University, 80 Waterman St, Box G-W, Providence, Rhode Island 02912, USA.

email: erika_edwards@brown.edu

²Department of Ecology and Evolutionary Biology, Yale University, PO Box 208105, New Haven, Connecticut 06520, USA.

Received 10 February; accepted 9 March 2015.

1. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
2. Zanne, A. E. *et al.* Corrigendum: Three keys to the radiation of angiosperms into freezing environments. *Nature* **514**, 394 (2014).
3. Zanne, A. E. *et al.* Corrigendum: Three keys to the radiation of angiosperms into freezing environments. *Nature* <http://dx.doi.org/10.1038/nature14371> (this issue).
4. Guralnick, R. P., Hill, A. W. & Lane, M. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecol. Lett.* **10**, 663–672 (2007).
5. Ponder, W. F., Carter, G. A., Flemons, P. & Chapman, R. R. Evaluation of museum collection data for use in biodiversity assessment. *Conserv. Biol.* **15**, 648–657 (2001).
6. Donoghue, M. J. & Edwards, E. J. Biome shifts and niche evolution in plants. *Annu. Rev. Ecol. Evol. Syst.* **45**, 547–572 (2014).

Author Contributions E.J.E., J.M.d.V. and M.J.D. designed the study; J.M.d.V. wrote all scripts for analyses; E.J.E. and J.M.d.V. analysed the data; E.J.E., J.M.d.V. and M.J.D. wrote the manuscript.

Competing Financial Interests Declared none.

doi:10.1038/nature14393

Zanne *et al.* reply

REPLYING TO E. J. Edwards, J. M. de Vos & M. J. Donoghue *Nature* **521**, <http://dx.doi.org/10.1038/nature14393> (2015)

Our goal was to understand which traits facilitated angiosperm shifts into freezing climates. Building on previous work^{1–5}, we showed strong support for evolutionary shifts in herbaceous habit, deciduous leaf phenology and small water-conducting conduits with the transition to exposure to freezing for the first time at this scale. We then decoupled the order of these shifts (traits-first versus climate-first pathways) based on a new summary of a long-standing method⁶ with no a priori expectations. Because current data sets are small compared to estimates of angiosperm diversity, our pathways analyses are preliminary. By establishing testable hypotheses and making our considerable resources public, future studies can build upon our questions. We found their suggestion in ref. 7 that we looked for reifying patterns in nature surprising. In the accompanying Comment⁸, Edwards *et al.*⁸ reanalysed data from Zanne *et al.*⁹, including removing data points and using new thresholds (below). After correcting an error in conduit size threshold¹⁰, we still found that "trait first" was the most likely pathway, albeit with less strength. Otherwise, we stand by the validity of our approaches.

Ancestral state estimates are notoriously unreliable¹¹. Rather than using estimates at hundreds or thousands of nodes, we used the presumably more reliable, inferred 8–12 transition rates to examine likely pathways. If the rate of going from state A to B is three times the rate of going from A to C for 100 species starting in state A, we expect 75 to go to B first and 25 to go to C first. This expectation follows directly from a summary of the rates. Calculations are more complex for four states, but result in the same information: converting rates to expected paths. Edwards *et al.*⁸ performed simulations that showed this summary was

not biased, but that known paths may deviate, at times substantially, from this expectation, especially if rates are similar.

We agree that various thresholds are potentially suitable⁸. We disagree that radically changing thresholds should reveal the same result; Edwards *et al.* varied cut-offs from requiring 0% to 100% of a species range experiencing freezing (see figure 1 in ref. 8) for that species to be freezing exposed. A priori, we targeted >2.5% of a species range. Edwards *et al.*⁸ targeted >50% of a species range. Both are valid and selection should be guided by the biology of the system. Under our definition, if a species experienced freezing somewhere, it had the potential to handle freezing (a species-specific trait, not unlike our leaf and stem traits). Owing to the limitations of GBIF coverage, we believe it was better to define species as experiencing freezing (with >2.5% allowing for outliers) rather than to expect a set amount of a species range to be in freezing.

We agree that narrowly defined case studies provide detailed insights into a given lineage⁸. Equally important, large-scale analyses afford synthesis, examining broad evolutionary hypotheses missed by narrow studies. These approaches are certainly complementary, each with strengths and weaknesses, and it is critical that studies continue to be conducted across multiple scales.

Amy E. Zanne^{1,2}, David C. Tank^{3,4}, William K. Cornwell^{5,6}, Jonathan M. Eastman^{3,4}, Stephen A. Smith⁷, Richard G. FitzJohn^{8,9}, Daniel J. McGlenn¹⁰, Brian C. O'Meara¹¹, Angela T. Moles⁶, Peter B. Reich^{12,13}, Dana L. Royer¹⁴, Douglas E. Soltis^{15,16,17},

BRIEF COMMUNICATIONS ARISING

Peter F. Stevens¹⁸, Mark Westoby⁹, Ian J. Wright⁹, Lonnie Aarssen¹⁹, Robert I. Bertin²⁰, Andre Calaminus¹⁵, Rafaël Govaerts²¹, Frank Hemmings⁶, Michelle R. Leishman⁹, Jacek Oleksyn^{12,22}, Pamela S. Soltis^{16,17}, Nathan G. Swenson²³, Laura Warman^{6,24} & Jeremy M. Beaulieu²⁵

¹Department of Biological Sciences, George Washington University, Washington DC 20052, USA.

²Center for Conservation and Sustainable Development, Missouri Botanical Garden, St Louis, Missouri 63121, USA.
email: aezanne@gmail.com

³Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844, USA.

⁴Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844, USA.

⁵Department of Ecological Sciences, Systems Ecology, de Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.

⁶Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia.

⁷Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA.

⁸Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia V6T1Z4, Canada.

⁹Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia.

¹⁰Department of Biology, College of Charleston, Charleston, South Carolina 29424, USA.

¹¹Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996, USA.

¹²Department of Forest Resources, University of Minnesota, St Paul, Minnesota 55108, USA.

¹³Hawkesbury Institute for the Environment, University of Western Sydney, Penrith, New South Wales 2751, Australia.

¹⁴Department of Earth and Environmental Sciences, Wesleyan University, Middletown, Connecticut 06459, USA.

¹⁵Department of Biology, University of Florida, Gainesville, Florida 32611, USA.

¹⁶Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, USA.

¹⁷Genetics Institute, University of Florida, Gainesville, Florida 32611, USA.

¹⁸Department of Biology, University of Missouri—St Louis, St Louis, Missouri 63121, USA.

¹⁹Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada.

²⁰Department of Biology, College of the Holy Cross, Worcester, Massachusetts 01610, USA.

²¹Royal Botanic Gardens, Kew, Richmond TW9 3AB, UK.

²²Polish Academy of Sciences, Institute of Dendrology, 62-035 Kornik, Poland.

²³Department of Plant Biology and Ecology, Evolutionary Biology and Behavior, Program, Michigan State University, East Lansing, Michigan 48824, USA.

²⁴Institute of Pacific Islands Forestry, USDA Forest Service, Hilo, Hawaii 96720, USA.

²⁵National Institute for Mathematical & Biological Synthesis, University of Tennessee, Knoxville, Tennessee 37996, USA.

1. Sinnott, E. W. & Bailey, I. W. The evolution of herbaceous plants and its bearing on certain problems of geology and climatology. *J. Geol.* **23**, 289–306 (1915).
2. Judd, W. S., Sanders, R. W. & Donoghue, M. J. Angiosperm family pairs: preliminary phylogenetic analysis. *Harv. Pap. Bot.* **1**, 5 (1994).
3. Moles, A. T. *et al.* Global patterns in plant height. *J. Ecol.* **97**, 923–932 (2009).
4. Wheeler, E. A., Baas, P. & Rodgers, S. Variations in dicot wood anatomy: a global analysis based on the InsideWood database. *IAWA J.* **28**, 229–258 (2007).
5. Botta, A., Viovy, N., Ciais, P., Friedlingstein, P. & Monfray, P. A global prognostic scheme of leaf onset using satellite data. *Glob. Change Biol.* **6**, 709–725 (2000).
6. Pagel, M. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. B* **255**, 37–45 (1994).
7. Edwards, E., de Vos, J. & Donoghue, M. The mega-fication of comparative biology. In *Botany 2014, Boise, Idaho* http://schenk.tulane.edu/Botany/ComPhy_files/botany2014_edwards.pdf.
8. Edwards, E., de Vos, J. M. & Donoghue, M. Doubtful pathways to cold tolerance in plants. *Nature* <http://dx.doi.org/10.1038/nature14393> (2015).
9. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92; corrected *Nature* **514**, 394 (2014).
10. Zanne, A. E. *et al.* Corrigendum: Three keys to the radiation of angiosperms into freezing environments. *Nature* <http://dx.doi.org/10.1038/nature14371> (this issue).
11. Cunningham, C. W., Omland, K. E. & Oakley, T. H. Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol. Evol.* **13**, 361–366 (1998).

doi:10.1038/nature14394

ARCHAEOLOGY

Tools go back in time

The finding of 3.3-million-year-old stone flints, cores, hammers and anvils in Kenya suggests that the first stone tools were made by human ancestors that pre-dated the earliest known members of the genus *Homo*. [SEE ARTICLE P.310](#)

ERELLA HOVERS

The earliest production of stone tools by human ancestors marked the dawn of an innovative behavioural strategy that transformed the ecology, social systems and culture of humans through social learning of skills and technology. Direct evidence for this behaviour is found in the archaeological record, and its earliest appearance has been gradually pushed deeper into time over the past five decades¹. On page 310 of this issue, Harmand *et al.*² report findings from Lomekwi 3 (LOM3), a site on the western side of Lake Turkana, Kenya, that extend the record of material culture by some 700,000 years, to 3.3 million years ago.

Palaeoanthropologists have long predicted, on evolutionary and archaeological grounds, that the first stone tools should be older than the previously oldest known instances, which date to around 2.6 million years ago (Ma). Once thought to be unique behaviours of hominins (the group that includes humans and their close extinct ancestors), tool use and toolmaking are now well documented in extant taxa of non-human primates — mainly chimpanzees, but also orangutans, gorillas and certain monkeys. Parsimony suggests that hominin tool use and toolmaking were practised by the last common ancestor of chimpanzees and hominins³, some 7–5 Ma, but that the making of tools from stones is unique to hominins, with the notable exception of chimpanzee nut-cracking hammer stones⁴; the tools of other primates are mostly twigs or other plant matter. However, the stone tools at sites dated to 2.6 Ma or slightly later, which are assigned to the archaeological complex known as the Oldowan¹, seem to be too well made to have been the first experiments of early humans in producing sharp-edged stone flakes by free-hand core-reduction techniques (when a stone block, or core, is held in the hand and hit with a hammer stone)^{5,6}.

Thus, core reduction as a technique for producing sharp edges for cutting plant and animal material was hypothesized to have first emerged among early hominin ancestors, but the time of its first appearance remained undefined in the absence of direct evidence. Furthermore, archaeologists did not have



Figure 1 | A Lomekwi tool. This stone was found on the surface at the Lomekwi 3 site in Kenya, where stone tools found in sedimentary layers have been dated to 3.3 million years old. The scars on the stone's surface indicate that it was used as a core from which flakes were produced.

a 'template' with which to search for these proposed very early stone tools, and open questions remained regarding the selective processes that might have led to the emergence of stone toolmaking.

Harmand and colleagues' findings inform us on the first two issues. Extensive geological mapping of the project area, and the presence of well-dated volcanic ash layers, together with palaeomagnetic correlations, constrained the age of the geological layer in which the authors had found the tools to 3.31–3.21 Ma. This range was narrowed down to 3.3 Ma on the basis of sedimentation rates. Because the sediments in these layers are fine-grained, and a flake found by the authors could be fitted back onto the core from which it had been detached, it is unlikely that the tools accumulated through stream activity or that substantial disturbance of the sediments occurred after the tools had been discarded.

The small collection of stone tools at LOM3 is unlike those from known Oldowan localities, which contain mainly flakes. Most of the LOM3 items (around 76%) are cores, anvils,

hammer stones and worked cobbles, indicating that the main activities were associated with pounding against anvils rather than hand-held core reduction. Some items bear traces of having been used for multiple techniques — as an anvil or hammer stone, or as a core to produce flakes — which hints at their use for exploiting variable resources (Fig. 1). All types of object in LOM3 are larger than their counterparts in the later Oldowan sites. The flakes are more massive than the accidentally produced flakes found in chimpanzee nut-cracking localities, but the anvils and hammers are within the size range of chimpanzee nut-cracking kits^{4,7,8}.

The surface characteristics of the items indicate the use of forceful blows during the application of both pounding and hand-held core reduction. However, whether the hand anatomy and precision grip that are needed for tool use and toolmaking were already present in the various species of hominin that existed around 3 Ma is debated^{9–11}. The earliest known¹² appearance of the genus *Homo* is at 2.8 Ma, suggesting that the LOM3 toolmakers may have belonged to other lineages of early hominin.

Harmand and colleagues contrast the postulated multiple uses for the LOM3 items with the generally single-purpose tools used by extant non-human primates. The authors suggest that the LOM3 tools could represent a technological stage between a hypothetical pounding-oriented stone-tool use by hominins earlier than those at LOM3 and the flaking-oriented behaviour of later Oldowan tool-makers. Primatologists may take issue with the first statement and argue that tool use in primates is multifaceted¹³. Evolutionary theorists may prefer less gradualist interpretations, and archaeologists could argue that one must not exclude the possibility that, at the beginning of each discrete episode of its use, each stone object was perceived merely as available raw material. The cognitive implications in this last case would differ from those offered by Harmand and colleagues. Therefore, we should focus on the evidence^{14,15} for core reduction as a marker of new cognitive abilities and a new technological path on which hominins embarked at 3.3 Ma.

The age and nature of the finds from LOM3 call for a re-evaluation of models^{16,17} that tie together the timing and patterns of environmental change, hominin evolution and the origins of technological behaviour around 2.5 Ma. However, caution is warranted. Our understanding of ancient hominins and their cognitive, cultural and social capacities is only as good as the available archaeological and fossil data. Similar to animal bones from Dikika, Ethiopia, that date to at least 3.39 Ma and arguably bear stone-inflicted cut marks¹⁸, the stone tools from LOM3 are at present an isolated occurrence. To maintain that either of these instances marks innovations in hominin behavioural evolution, the temporal gaps must be filled in with more data. In this respect, the LOM3 discoveries stand to have an immediate impact on human-origins research in eastern Africa by providing the long-needed search template for early stone tools.

Moreover, until now, the search for 'older than the Oldowan' archaeological sites has focused on a few areas that contain sediments dated to between 2.9 and 2.6 Ma, with the aim of establishing a sequence with known Oldowan sites. The discoveries at LOM3 allow research also to focus on the time range 3.4–2.9 Ma, which so far has not been tapped for evidence of material culture. And why not dig deeper in time? LOM3 may not be the final — or rather, the first — word on the roots of human technology. ■

Erella Hovers is at the Institute of Archaeology, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel.
e-mail: hovers@mail.huji.ac.il

- de la Torre, I. *Phil. Trans. R. Soc. B* **366**, 1028–1037 (2011).
- Harmand, S. *et al. Nature* **521**, 310–313 (2015).
- Panger, M. A., Brooks, A. S., Richmond, B. G. &

- Wood, B. *Evol. Anthropol.* **11**, 235–245 (2002).
- Carvalho, S., Cunha, E., Sousa, C. & Matsuzawa, T. *J. Hum. Evol.* **55**, 148–163 (2008).
- Semaw, S. *et al. J. Hum. Evol.* **45**, 169–177 (2003).
- Delagnes, A. & Roche, H. *J. Hum. Evol.* **48**, 435–472 (2005).
- Mercader, J., Panger, M. & Boesch, C. *Science* **296**, 1452–1455 (2002).
- Mercader, J. *et al. Proc. Natl Acad. Sci. USA* **104**, 3043–3048 (2007).
- Marzke, M. W. *Phil. Trans. R. Soc. B* **368**, 0120414 (2013).
- Rolian, C. & Gordon, A. D. *Am. J. Phys. Anthropol.*

- 152**, 393–406 (2013).
- Skinner, M. M. *et al. Science* **347**, 395–399 (2015).
- Villmoare, B. *et al. Science* **347**, 1352–1355 (2015).
- McGrew, W. C. *Science* **328**, 579–580 (2010).
- Read, D. & van der Leeuw, S. *Phil. Trans. R. Soc. B* **363**, 1959–1968 (2008).
- Hovers, E. in *Origins of Human Creativity and Innovation Developments in Quaternary Science* Vol. 16 (ed. Elias, S.) 51–68 (Elsevier, 2012).
- Vrba, E. S. *Nat. Hist.* **102**, 47–51 (1993).
- Potts, R. *Curr. Anthropol.* **53**, S299–S317 (2012).
- McPherron, S. P. *et al. Nature* **466**, 857–860 (2010).

QUANTUM PHYSICS

Squeezed ions in two places at once

Experiments on a trapped calcium ion have again exposed the strange nature of quantum phenomena, and could pave the way for sensitive techniques to explore the boundary between the quantum and classical worlds. **SEE LETTER P.336**

TRACY NORTUP

In Schrödinger's famous thought experiment, a cat is prepared in a quantum superposition of being both alive and dead by being trapped in a box with a flask of poison. As if that were not enough, the poor cat is now being squeezed too — all in the name of quantum measurement. In laboratory experiments, atoms have been prepared in superpositions of being in two places at once, playfully called Schrödinger's cat states¹. On page 336 of this issue, Lo *et al.*² demonstrate superposition states of a trapped ion in which its position is not only split between two locations, but also squeezed. Squeezing refers to the process of suppressing quantum fluctuations for a particular measurement, such as that of a particle's position.

Quantum mechanics tells us that the

position of a particle (or Schrödinger's fictitious cat) has an inherent uncertainty even when it is at rest, a feature known as the standard quantum limit. When the particle is prepared in a squeezed state, however, we can pinpoint its position to better than that limit (Fig. 1). There is a price to pay for squeezing, though. When fluctuations in position are squashed down, additional fluctuations arise in the particle's momentum, such that the product of position and momentum fluctuations still satisfies Heisenberg's uncertainty relation — which states that there is a fundamental limit to the precision with which a particle's position and momentum can be simultaneously determined. Nevertheless, by suppressing fluctuations in the quantity that they intend to measure, researchers can improve measurement precision. For example, squeezed states have been used to achieve record

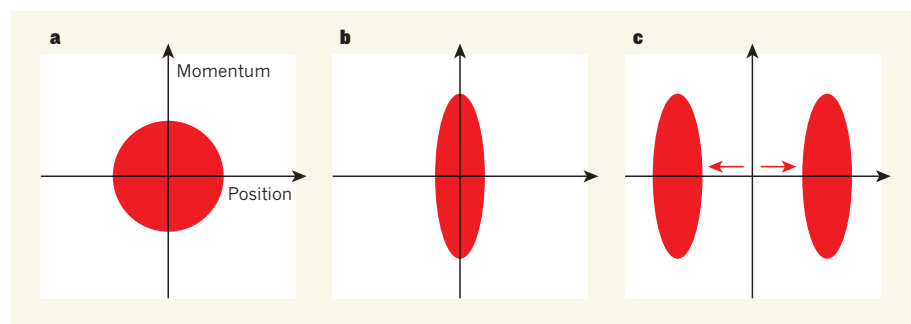


Figure 1 | Squeezing an ion's positional uncertainty. Every object's momentum and position are subject to fluctuations, which become pronounced on the atomic scale. **a**, The red circle indicates the uncertainty in position and momentum for a calcium ion (Ca^+) in its motional ground state. **b**, Lo *et al.*² used laser pulses to squeeze fluctuations in position, at the cost of amplifying the fluctuations in momentum. **c**, They then displaced the ion in opposite directions at once, so that it would be equally likely to be found in one of two distinct states. The squeezing operation provides a better signal-to-noise ratio for the ion's position, so that it is easier to distinguish between the states.

sensitivities for one of the detectors at the Laser Interferometer Gravitational-Wave Observatory in Richland, Washington³.

The starting point for Lo and colleagues' study is a single calcium ion (Ca^+) trapped by radiofrequency electromagnetic fields in a vacuum vessel. One can picture the trapped ion as a tiny pendulum oscillating around its equilibrium position. For a quantum pendulum in its lowest energy state, the uncertainties in its position and momentum have equal magnitude. In this case, squeezing corresponds to suppressing position fluctuations at the cost of momentum, or vice versa.

The authors use a set of methods known as laser cooling to bring the ion to its motional ground state⁴, and then introduce additional laser fields to squeeze the state, reducing the positional variance by a factor of nine. Although squeezed states of trapped ions were first demonstrated 19 years ago⁵, the fidelity with which these delicate states are prepared is highly sensitive to experimental noise, such as fluctuating electric and magnetic fields. The authors used a technique called reservoir engineering, which was previously developed by the same research group⁶, to achieve robust, high-fidelity squeezing even in the presence of noise.

With the ion in a squeezed ground state, the next step is to prepare it in a cat-state superposition. Imagine that the ion pendulum is displaced by pulling it to one side, then releasing it; it will swing back and forth with the amplitude that has been imparted. Now imagine pulling the ion to the right and left at the same time: classically this does not make sense, but quantum mechanically it is possible.

The way to do this with a trapped ion is to apply a state-dependent force — a displacement whose direction depends on the spin state of the ion's outermost electron¹. When the electron is prepared in a superposition of two spin states, the force acts in an equal and opposite direction on each component. As a result, the ion pendulum's motion is a superposition of two possible oscillations, each with the same amplitude but in opposite directions. In fact, each motional direction is entangled with the electron's spin state; that is, one property cannot be described independently of the other.

How distinguishable are the two cat-state components from each other? It depends on whether the initial squeezing was performed on the ion's position or on its momentum. Lo and colleagues measured and compared the two cases. If momentum fluctuations were suppressed before the cat state was prepared, then the corresponding enhancement in position fluctuations made the spatial separation more difficult to distinguish. By contrast, if the ion's position was squeezed, then the spatial separation between the components became 56 times larger than the extent of the squeezed positional fluctuations.

It is exactly this amplified sensitivity to spatial separation that makes squeezed states promising for future applications. For example, using cat states, the wave nature of a single ion can be exploited for interferometry. In an interferometer, a wave is split, sent along two paths and finally recombined, providing information about how the paths differ. In a cat state, the ion's location is split into two superposition components, each of which explores a different path. Thus, if the cat-state components are recombined, the superposition acts as an interferometer, probing path differences. Moreover, an ion is highly sensitive to changes in electric and magnetic fields, which shift its electron energy levels, so an ion interferometer could measure field gradients on the scale of tens of nanometres⁷. Squeezed cat states would also be more robust than non-squeezed states to certain types of noise, providing improved sensing capabilities.

Building on established techniques for the precise manipulation of trapped ions, the authors have demonstrated an exciting new capability for both engineering and characterizing quantum states. These states are fascinating, not only as future sensors, but also as a means of exploring the boundary between the quantum and classical worlds. The ion

pendulum demonstrated by Lo and colleagues has a position uncertainty of only a few nanometres, but it swings back and forth — in two directions at once — over hundreds of nanometres, a much larger distance than atomic scales. Efforts are under way in many research groups to extend cat-state length scales even further, into truly macroscopic regimes.

Future work with squeezed cat states will continue to characterize their strange, often counter-intuitive, quantum properties. Here, as the authors have shown, single ions provide an exceptional experimental platform on which to do so. ■

Tracy Northup is at the *Institut für Experimentalphysik, Universität Innsbruck, Innsbruck 6020, Austria.*
e-mail: tracy.northup@uibk.ac.at

1. Monroe, C., Meekhof, D. M., King, B. E. & Wineland, D. J. *Science* **272**, 1131–1136 (1996).
2. Lo, H.-Y. *et al. Nature* **521**, 336–339 (2015).
3. Aasi, J. *et al. Nature Photon.* **7**, 613–619 (2013).
4. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. *Rev. Mod. Phys.* **75**, 281–324 (2003).
5. Meekhof, D. M., Monroe, C., King, B. E., Itano, W. M. & Wineland, D. J. *Phys. Rev. Lett.* **76**, 1796–1799 (1996).
6. Kienzler, D. *et al. Science* **347**, 53–56 (2015).
7. Poyatos, J. F., Cirac, J. I., Blatt, R. & Zoller, P. *Phys. Rev. A* **54**, 1532–1540 (1996).

STEM CELLS

Asymmetric rejuvenation

Organelles called mitochondria are asymmetrically apportioned to the daughters of dividing stem cells according to mitochondrial age. This finding sheds light on the mechanisms underlying asymmetric stem-cell division.

ANU SUOMALAINEN

The thought of reversing the ageing process has tickled the human imagination for centuries. Despite the air of mystery surrounding the topic, rejuvenation occurs so naturally that we pay no attention to it — that is, when mothers give birth to offspring. Although babies originate from the germ cells of a mother and father who might be decades old, they do not inherit their parents' accumulated cellular damage, but get a fresh start. Writing in *Science*, Katajisto *et al.*¹ suggest that such rejuvenation may also be a characteristic of the stem cells responsible for tissue maintenance.

Stem cells have some distinctive characteristics. They are long-lived, or even immortal, and can divide asymmetrically². The difference between the daughter cells of an asymmetric stem-cell division is not subtle. One

daughter inherits the mother's immortality and ability to give rise to many cell types. The other must leave the cosy stem-cell home, become mortal and commit to differentiating into a cell with a specialized identity, for example a cell of the gut wall, eschewing its broad potential in favour of excelling at one particular task.

Katajisto *et al.*¹ focused on the stem cells of human mammary tissue. Samples taken from the tissue and cultured *in vitro* contain small, round, stem-cell-like cells and flat epithelial cells, which line the mammary ducts *in vivo*. The different daughters of mammary stem-cell divisions are therefore easily distinguished by microscopy, and their fates can be followed *in vitro*. To investigate whether asymmetric stem-cell division involves asymmetric apportioning of organelles to the two daughters, the authors developed assays that enabled them to tag organelles and then activate the tags at

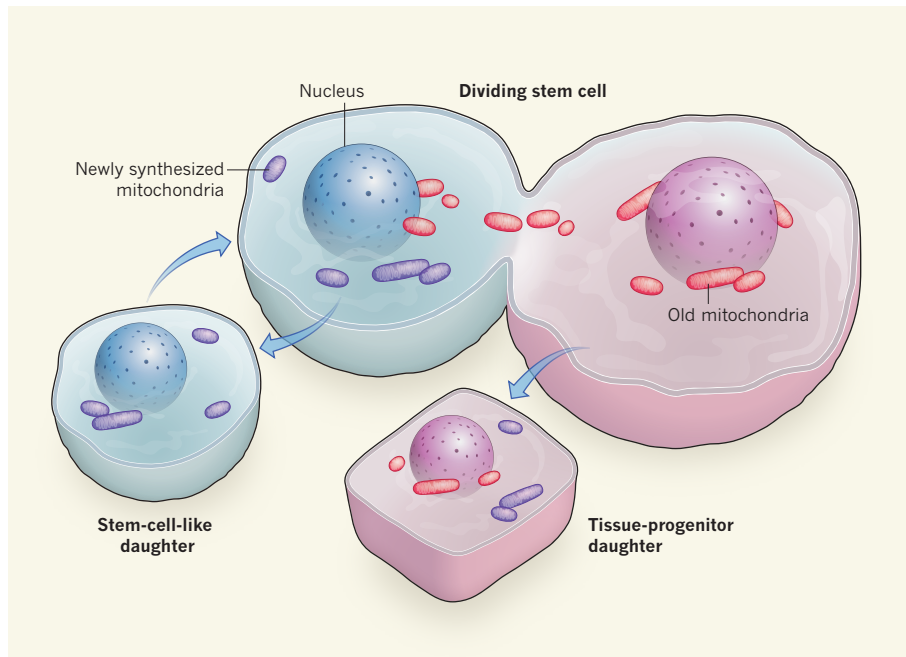


Figure 1 | Unequal sharing between daughters. Tissue stem cells undergo asymmetric cell division, in which one daughter cell adopts a stem-cell-like state and the other differentiates into a more-specialized cell type. Katajisto *et al.*¹ report that organelles called mitochondria are split unevenly between the two daughters. Older organelles, which are located in the region surrounding the nucleus of the mother cell, are apportioned primarily to the tissue-progenitor daughter, whereas newly synthesized mitochondria are apportioned to the stem-cell-like daughter.

exact times. In this way, they could identify the organelles that were newly synthesized and those that were old, and track them after cell division.

The researchers observed that the various types of organelle were similarly distributed between the two daughters, with one exception. Organelles called mitochondria showed differential segregation, such that the multi-talented stem-cell daughter received most of the newly synthesized mitochondria, whereas the tissue-progenitor daughter received around six times more old mitochondria (Fig. 1). Thus, organellar rejuvenation occurs in tissue stem cells, and involves mitochondria.

Mitochondria use oxygen to burn fats, sugars and amino acids, generating ATP molecules that act as the cell's energy currency. This oxidative metabolism establishes an electric charge (a membrane potential) across the membrane surrounding the organelle³ that can be used as a measure of mitochondrial ATP synthesis. Oxidative metabolism also generates side products in the form of reactive oxygen species (ROS) — potent signalling molecules that, if produced in excess, can damage surrounding proteins, lipids and DNA. Subtle changes in ROS can modify stem-cell behaviour, promoting commitment to differentiation⁴. Indeed, mitochondrial dysfunction promotes stem-cell dysfunction and exhaustion, leading to premature signs of ageing that mimic physiological ageing^{5–7}. By contrast, fully functional mitochondrial proteins minimize ROS production and maximize control

over oxidative metabolism. It is therefore no surprise that stem cells treasure prime fitness in this organelle.

The concept of apportioning old mitochondria asymmetrically has already been established in baker's yeast, in which damaged proteins and mitochondria with lower oxidative function preferentially remain in the mother cell^{8,9}, rather than entering the daughter that buds off from it. By contrast, Katajisto *et al.* found no functional differences between the mitochondria apportioned to the different daughter cells, because the membrane potential was similar in both types of cell. Even when the authors abolished the membrane potential, asymmetric apportioning occurred. In fact, the only determinant of mitochondrial fate was organellar age.

Katajisto and colleagues showed that ageing mitochondria were preferentially located close to the nucleus, whereas young organelles were also found in the cell periphery (Fig. 1). This suggests that physical segregation of the organelles contributes to differential delivery during cell division. Chemical inhibition of the fission process by which mitochondria divide hindered this compartmentalization, indicating a key role for mitochondrial dynamics in asymmetric mitochondrial segregation.

Asymmetric mitochondrial apportioning could be an indication of the general selfishness of stem cells — the cells that end up being mortal are largely unimportant compared with their immortal sisters. This hypothesis would

be consistent with the 'disposable soma' theory of ageing¹⁰ (extended here to apply to tissue stem cells), which posits that an organism is merely disposable packing material for its germ cells. The second possibility, however, is that the committed daughter cell actually requires old mitochondria to fulfil its function. Mitochondrial ATP synthesis increases on differentiation, and an increase in ROS in response to increased mitochondrial function is associated with differentiation⁴. For example, in red blood cells, subtle increases in ROS orchestrate iron loading and cell maturation¹¹. The asymmetric apportioning of mitochondria could therefore provide the ROS boost required to initiate a differentiation program.

The ultimate fate of old mitochondria during the differentiation of tissue-progenitor daughters remains an open question. Eventually they will be recycled, and new organelles will replace them. The authors noticed that asymmetric apportioning of mitochondria required the presence of parkin, a protein that marks mitochondria for recycling¹². However, there were no apparent changes in recycling levels in daughter cells. Whether parkin has a role, for example, in the timing of degradation of the old organelles after division remains unknown.

Katajisto and colleagues' study raises questions about the role of mitochondrial quality control as a regulator of cell fate and behaviour. For instance, an exciting possibility is that the mechanism described is a general feature of stem cells. It will be interesting to investigate whether similar mechanisms are in place in mature tissues. Furthermore, it is unclear how stem cells would handle increased mitochondrial-protein damage in mitochondrial disorders.

Another avenue for study is what happens to mitochondrial DNA during asymmetric mitochondrial apportioning. And finally, do similar mechanisms apply in germ cells, providing the offspring with a fresh mitochondrial start? Defining the molecular mechanisms underlying this phenomenon will bring us a step closer to understanding the cellular recipe for immortality — the rejuvenation of energy metabolism. ■

Anu Suomalainen is in the Research Programs Unit, Molecular Neurology, Biomedicum-Helsinki, University of Helsinki, Helsinki 00290, Finland.
e-mail: anu.wartiovaara@helsinki.fi

1. Katajisto, P. *et al.* *Science* **348**, 340–343 (2015).
2. Mukherjee, S., Kong, J. & Brat, D. J. *Stem Cells Dev.* **24**, 405–416 (2015).
3. Nunnari, J. & Suomalainen, A. *Cell* **148**, 1145–1159 (2012).
4. Ito, K. & Suda, T. *Nature Rev. Mol. Cell Biol.* **15**, 243–256 (2014).
5. Ahlqvist, K. J. *et al.* *Cell Metab.* **15**, 100–109 (2012).
6. Trifunovic, A. *et al.* *Nature* **429**, 417–423 (2004).

7. Kujath, G. C. *et al. Science* **309**, 481–484 (2005).
 8. Aguilaniu, H., Gustafsson, L., Rigoulet, M. & Nyström, T. *Science* **299**, 1751–1753 (2003).

9. McFaline-Figueroa, J. R. *et al. Aging Cell* **10**, 885–895 (2011).
 10. Kirkwood, T. B. L. *Nature* **270**, 301–304 (1977).

11. Ahlqvist, K. J. *et al. Nature Commun.* **6**, 6494 (2015).
 12. Narendra, D., Tanaka, A., Suen, D.-F. & Youle, R. J. *J. Cell Biol.* **183**, 795–803 (2008).

MATERIALS SCIENCE

Magnetic alloys break the rules

A family of alloys has been discovered that undergoes unexpected changes of shape when magnetized. This strange behaviour might help in unravelling the mystery of a phenomenon called magnetic hysteresis. [SEE LETTER P.340](#)

RICHARD D. JAMES

One of the biggest puzzles in materials science is magnetic hysteresis — an effect in which the magnetization of a material depends on present and past applied magnetic fields. Hard magnets, which resist demagnetization in magnetic fields, have exceptionally large hysteresis (Fig. 1a). By contrast, soft magnets become fully magnetized under a small applied magnetic field, spontaneously demagnetize when the field is removed and have the smallest known hysteresis (Fig. 1b). But so little is known about hysteresis that rule-of-thumb strategies for developing alloys intended to increase the hardness of magnets sometimes lead to softer ones. The report by Chopra and Wuttig¹ (page 340) of a new kind of magnetic alloy that has near-zero hysteresis presents fresh opportunities to study the mechanism of this perplexing phenomenon.

The need to understand magnetic hysteresis is becoming increasingly urgent^{2,3}. That is

because hard magnets are main components of the motors of electric vehicles and technologies such as wind-power generators, both of which are becoming more widespread. Soft magnets are ubiquitous in the electronic devices that control power in electric motors and current flow in electrical grids.

So why is hysteresis so hard to comprehend? It depends intimately on the behaviour of magnetic domains in a material^{4,5}, which can be exceedingly complex, as Chopra and Wuttig describe. But despite the geometric complexity, domain configuration is ultimately determined by a few fundamental material constants and by the shape of the magnetic body. Magnetic domains may also interact with the boundaries of the small crystals (grains) from which the material is composed, and with defects — non-magnetic impurities that inevitably form in alloys — in ways that affect hysteresis. How the fundamental constants conspire with these structural aspects to deliver a specific hysteresis profile is little understood. But hysteresis profiles (loops)

are quite reproducible in alloys that have the same composition and that are processed in similar ways.

One of the constants mentioned above, K_1 , quantifies the difficulty of rotating the direction of magnetization within a crystal. Most researchers regard it as the most important fundamental property affecting hysteresis: a high K_1 is associated with large hysteresis, and a low K_1 with small hysteresis. The constants λ_{100} and λ_{111} are also thought to be relevant. These quantify magnetostriction (changes in shape of materials owing to magnetization) in two crystallographic directions, defined as 100 and 111; high values indicate large shape changes.

But the precise role of these fundamental constants is not clear. For example, two of the softest magnetic materials are permalloys⁶: one is 55% iron and 45% nickel; the other contains 21.5% iron and 78.5% nickel. The K_1 for the first permalloy is quite large, which suggests that the alloy should have a large hysteresis. But its hysteresis is actually small, so K_1 alone does not tell the whole story. Moreover, K_1 becomes zero for an iron–nickel mixture that has 75% nickel, which suggests that this alloy should be particularly soft — but it is not. By contrast, the (softest) 78.5% nickel permalloy has a fairly small, non-zero K_1 , which suggests that it should be harder than the 75% nickel alloy, but it is not.

Intriguingly, the magnetostriction constants also become zero at nickel compositions close to 78.5%, but not at precisely that composition: λ_{100} is zero at 83%, whereas λ_{111} is zero at 80%. The only other composition of iron–nickel alloys at which λ_{100} is zero is precisely that of the first permalloy (45% nickel), which suggests that magnetostriction is relevant to softness.

The discrepancies between the values of the constants and the resulting magnetic behaviour constitute the long-standing ‘permalloy problem’⁷. Although there are vague comments in the literature about the possible involvement of stress in determining hysteresis in these materials, its precise role is unclear. We certainly do not have a theory that predicts the composition of the softest permalloy to be 78.5% nickel.

To unravel this mystery, it is crucial to find magnetic materials that exhibit new behaviours. Chopra and Wuttig have discovered a family of alloys that behave in three unusual ways. First, the alloys are decidedly ‘non-Joulian’: most magnetic materials retain their overall volume during magnetostriction, whereas the authors’ alloys expand considerably. A major implication of this is that the model used almost universally to quantify

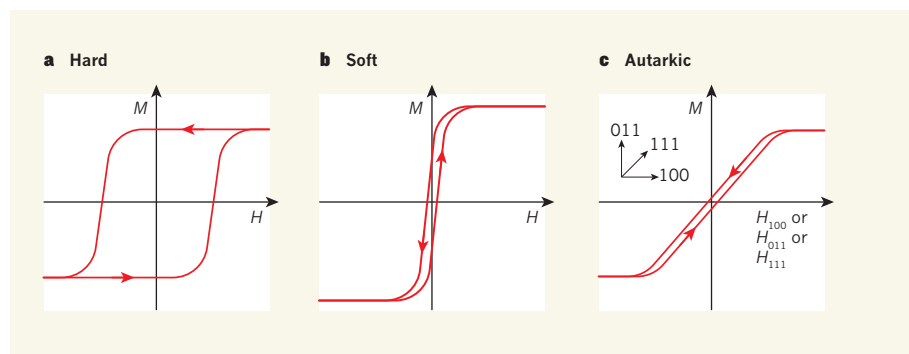


Figure 1 | Typical hysteresis loops in hard, soft and autarkic magnets. Hysteresis loops define the average magnetization, M , of a material in the direction of an applied alternating magnetic field, H , as the field is increased and then decreased. **a**, Hard magnets are difficult to magnetize and demagnetize, and have the largest hysteresis loops. Arrows indicate whether H was increasing or decreasing as each part of the loop was measured. **b**, Soft magnets become fully magnetized in a small magnetic field and exhibit small hysteresis. **c**, Chopra and Wuttig¹ report alloys with ‘autarkic’ magnetic domains that undergo large, direction-dependent shape changes in magnetic fields, and which require a fairly large field to be magnetized, but have near-zero hysteresis and the same magnetization curves in all directions of a crystal. The indices 011, 111 and 100 denote different directions in the crystal lattice of the material in which the magnetic field is applied.

magnetization-induced shape changes⁶ (which involves only the constants λ_{100} and λ_{111}) fails to describe such changes in the new alloys.

Second, the researchers' samples are quite difficult to magnetize, as seen from the modest slope of the plots of magnetization against applied magnetic field (Fig. 1c). Intriguingly, they are equally hard to magnetize in all directions of the crystal lattice, even though shape changes owing to magnetostriction are highly direction dependent. Third, and perhaps most interestingly, the materials are nearly hysteresis free.

Chopra and Wuttig explain these unexpected behaviours by proposing a concept called autarky: the independent action of magnetic domains⁵. These domains adopt the shape of a classic rectangular Landau pattern (see Fig. 3c of the paper¹) decorated by tiny zigzags, and are formed from several sub-regions. The domains undergo local, rather large shape changes during magnetization, but no domain seems to influence its neighbours. This behaviour is no doubt aided by the fact that the magnetization of each domain can point in any direction — that is, the magnetic force associated with a domain does not favour any particular direction.

To put it another way, in the absence of an applied field, it is as if each subregion within a rectangle is a piece of a jigsaw puzzle, but each piece is substantially distorted from its ordinary, regular shape. Nevertheless, the puzzle stays perfectly flat and rectangular rather than buckling up out of its plane, with each distorted piece fitting perfectly with its neighbours.

Given that further development of energy technology hinges largely on making exceptionally soft and exceptionally hard magnets, and that we understand so little about magnetic hysteresis, Chopra and Wuttig's finding is not only a dramatic fundamental discovery, but it could also be a touchstone on the way to a predictive theory of magnetic hysteresis. As a first step, it will be essential to find out in detail the changes of strain and magnetization that occur in the autarkic domains during magnetization. ■

Richard D. James is in the Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis 55455, USA.
e-mail: james@umn.edu

1. Chopra, H. D. & Wuttig, M. *Nature* **521**, 340–343 (2015).
2. Gutfleisch, O. *et al.* *Adv. Mater.* **23**, 821–842 (2011).
3. Coey, J. M. D. *IEEE Trans. Magn.* **47**, 4671–4681 (2011).
4. Bertotti, G. & Mayergoyz, I. D. *The Science of Hysteresis* (Gulf, 2006).
5. Hubert, A. & Schäfer, R. *Magnetic Domains: The Analysis of Magnetic Microstructures* (Springer, 2008).
6. Bozorth, R. M. *Ferromagnetism* (Van Nostrand, 1951).
7. Bozorth, R. M. *Rev. Mod. Phys.* **25**, 42 (1953).

STEM CELLS

Equilibrium established

Pluripotent cells can produce all cell types in the body. It emerges that this state of potential is endowed by cues, including inhibition of Wnt signalling, that maintain a balance between diverse cellular outcomes. [SEE ARTICLE P.316](#)

KYLE M. LOH & BING LIM

Achieving dualism — a state in which two opposing forces coexist in balance — is central to Taoist philosophy, and, it has now emerged, to stem cells too. Stem cells reside at a nexus of opportunity, harbouring the potential to form myriad tissues, from blood to bone to brain. Balancing these diverse potentials is key to endowing and maintaining stem-cell identity¹. On page 316 of this issue, Wu *et al.*² show that neutralizing one cellular signalling pathway, Wnt, helps stem cells to achieve such balance.

Stem cells that can form all bodily tissue types are said to be pluripotent³. Pluripotency is not a singular state, but is a property of at least two related developmental cell types. The first pluripotent cells to arise in mouse embryos have broad cellular potential and are dubbed naive⁴. Soon after naive cells form, they become primed for differentiation⁴, as many extracellular signals, including Fgf and Wnt proteins, direct them to become one of various specialized cell types. Specifically, primed pluripotent cells can become either ectoderm (the progenitor to skin and brain tissue) or mesendoderm (the progenitor

to blood, bone, intestines and other organs)³ (Fig. 1a).

Because primed pluripotent cells are poised to undergo imminent differentiation, they exist in a precarious position¹. If taken from an embryo and cultivated in a Petri dish, primed cells often spontaneously lose pluripotency, and develop into differentiated cell types⁵. This is partly attributable to the action of Wnt and Fgf proteins, which both induce mesendoderm differentiation and block ectoderm formation (Fig. 1a)^{6,7}.

Primed pluripotent cells produce Wnt, and might thereby intrinsically prompt their own differentiation^{5,8,9}. Wu *et al.* thus reasoned that they could block mesendoderm differentiation in this cell type by blocking Wnt^{5,10,11} and simultaneously restrict ectoderm formation by supplying Fgf (Fig. 1b). By stabilizing a seesaw of opposing lineage forces, an uncommitted pluripotent state might be realized at the fulcrum. The authors found that such treatment broadly 'stabilized' primed pluripotent cells, whether of human, macaque, chimpanzee or mouse provenance.

To investigate whether primed pluripotent cells stabilized in this manner retain the potential to develop into ectodermal and

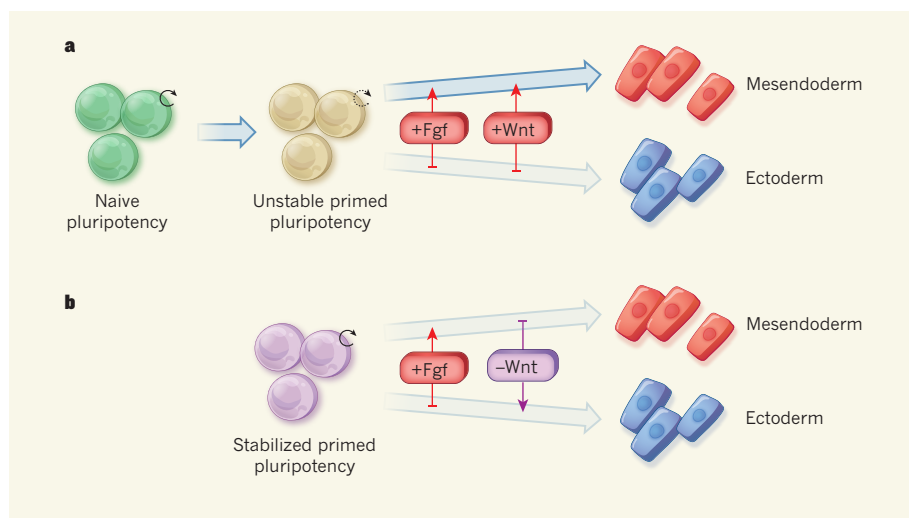


Figure 1 | Stabilizing the stem-cell seesaw. **a**, During normal development, naive pluripotent cells, which have the potential to give rise to all bodily cell types, mature into unstable primed pluripotent cells. These primed cells differentiate into more-specialized cell types (mesendoderm or ectoderm) in response to various signalling pathways. Fgf signalling and Wnt signalling both block ectoderm formation while promoting mesendoderm formation. **b**, Wu *et al.*² demonstrate that primed pluripotent cells are perched on a precarious seesaw between mesendoderm and ectoderm fates. By providing Fgf signals and simultaneously inhibiting Wnt signals, primed pluripotent cells can be stabilized.

mesendodermal cells, Wu *et al.* grafted stabilized human pluripotent stem cells onto 7.5-day-old post-implantation mouse epiblasts placed in a Petri dish (epiblasts are isolated, non-intact embryonic tissue fragments that lack supporting tissues and are therefore not viable). Strikingly, the stabilized human pluripotent cells successfully integrated into these mouse epiblasts, and the engrafted cells seemed to resume their natural developmental programme, differentiating into cells that expressed human ectoderm- and mesendoderm-specific genes in the confines of the epiblast. Although the full repertoire of the developmental genes expressed awaits a more extensive analysis, these findings imply that stabilized pluripotent cells are still capable of differentiation once released from stabilizing conditions.

Do the stabilized pluripotent cells correspond to any natural cellular state on the timeline of *in vivo* development? The classification of pluripotent cells as either naive or primed is probably an artificial dichotomy, and, indeed, gene expression in the Wnt-inhibitor-grown cells differs from that of either naive or primed pluripotent cells. Does this mean that such stabilized cells are genuinely a different class of pluripotent cell, or do they simply represent a more stabilized type of primed pluripotency, owing to a rebalancing of competing lineage forces? Perhaps 'stabilized' primed pluripotency is short-lived *in vivo* because of the speed of embryonic development, complicating efforts to assign *in vivo* counterparts to these cells. Some evidence¹⁰ argues that the stabilized cells correspond to an intermediate between naive and primed pluripotency.

A final possibility is that Wu and colleagues' cells exist orthogonally to the natural developmental timeline — that is, they are an artificial, non-developmental cell type. Maybe the priming of these cells has not been rewritten by Wnt inhibition at all. Instead, a change in adhesion properties could enable the stabilized human cells to engraft into the isolated mouse epiblast *in vitro*. Perhaps reflecting some degree of artificiality, the stabilized cells engraft only into the posterior of such epiblasts, whereas conventional primed cells from mice can engraft into all regions. This bias remains unexplained.

Finally, we propose that the idea of lineage balance¹ might not be specific to pluripotent stem cells, but might also extend to more-specialized ones, such as gut¹² or blood¹³ stem cells. If stem cells represent a state in which opposing lineage potentials coexist, then negotiating a balance in competing lineage forces might prove decisive in stabilizing and thus capturing diverse types of stem cell. ■

Kyle M. Loh is in the Department of Developmental Biology and at the Stanford Institute for Stem Cell Biology & Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA.

Bing Lim is at Merck Research Laboratories, Translational Medicine Research Centre, Singapore 138665.
e-mail: kyleloh@stanford.edu;
bing.lim@merck.com

1. Loh, K. M. & Lim, B. *Cell Stem Cell* **8**, 363–369 (2011).
2. Wu, J. *et al.* *Nature* **521**, 316–321 (2015).
3. Loh, K. M., Lim, B. & Ang, L. T. *Physiol. Rev.* **95**, 245–295 (2015).
4. Nichols, J. & Smith, A. *Cell Stem Cell* **4**, 487–492 (2009).

5. Tsakiridis, A. *et al.* *Development* **141**, 1209–1221 (2014).
6. Greber, B. *et al.* *EMBO J.* **30**, 4874–4884 (2011).
7. Loh, K. M. *et al.* *Cell Stem Cell* **14**, 237–252 (2014).
8. Blauwkamp, T. A., Nigam, S., Ardehali, R., Weissman, I. L. & Nusse, R. *Nature Commun.* **3**, 1070 (2012).
9. Sugimoto, M. *et al.* *Stem Cell Rep.* **4**, 744–757 (2015).
10. Kurek, D. *et al.* *Stem Cell Rep.* **4**, 114–128 (2015).
11. Sumi, T., Oki, S., Kitajima, K. & Meno, C. *PLoS ONE* **8**, e63378 (2013).
12. Yin, X. *et al.* *Nature Meth.* **11**, 106–112 (2014).
13. Fares, I. *et al.* *Science* **345**, 1509–1512 (2014).

MOLECULAR BIOLOGY

Splicing does the two-step

An intricate recursive RNA splicing mechanism that removes especially long introns (non-coding sequences) from genes has been found to be evolutionarily conserved and more prevalent than previously thought. SEE LETTERS P.371 & P.376

**HEIDI COOK-ANDERSEN
& MILES F. WILKINSON**

One of the biggest surprises in molecular biology was the discovery in 1977 that coding information in genes is interrupted by non-coding sequences known as introns. Much has since been learned about how introns are recognized and spliced out of precursor RNA to yield mature messenger RNA in which the remaining sequences — the exons — are stitched together. A lingering challenge has been to work out the way in which long introns are correctly recognized and spliced out, because they have a greater potential for splicing errors than do short introns.

One intriguing solution to this problem arrived 17 years ago, with the discovery that a long intron in the *Ultrabithorax* gene in the fruit fly *Drosophila melanogaster* is removed in a progressive, stepwise fashion, thereby reducing the size of the chunks that need to be defined for splicing¹. However, subsequent studies identified only a handful of fly genes that undergo this 'recursive' splicing^{2,3}, and no examples were demonstrated in other species⁴, casting doubt on the generality of the process. Two papers in this issue report that recursive splicing is actually quite widespread in fly genes⁵ and that it is also used by genes expressed in the human brain^{5,6}.

Recursive splicing depends on juxtaposed 3' and 5' splice-site sequences, called recursive splice sites, in the middle of long introns (Fig. 1a). Duff *et al.*⁵ (page 376) set out to identify recursive splice sites in *D. melanogaster* using deep-sequencing methods. Their screen yielded 197 functional recursive splice sites, many of which were highly conserved across several *Drosophila* strains. The authors

identified a total of 115 fly genes that undergo recursive splicing, greatly expanding the range of this mechanism.

By evaluating the spliced-out intron segments (lariats), Duff *et al.* obtained evidence that recursive splicing is a sequential and largely obligate process for genes that have recursive splice sites. They also found that recursive 3' splice sites are typically richer in the long tracts of pyrimidines (the nucleotide bases cytosine and uracil) required for splicing than are non-recursive 3' splice sites. This raises the possibility that their splicing depends more than that of typical introns on the polypyrimidine-tract-binding protein U2AF. Indeed, the authors found that recursive splicing is strikingly more sensitive to U2AF depletion than is canonical splicing. The physiological significance of this intriguing discovery remains to be determined.

Sibley *et al.*⁶ (page 371) addressed the long-standing question of whether recursive splicing is evolutionarily conserved. Using two complementary approaches, they identified nine genes that undergo recursive splicing in the human brain. In contrast to sites in *Drosophila*, in which the majority of recursive introns are completely spliced out^{1–3,5}, all recursive splice sites identified in humans harboured an 'RS exon' that seems to be pivotal for removing the long intron and can be retained in some circumstances (Fig. 1b).

The authors identified two roles for the RS exon in recursive splicing in humans. First, it facilitates recognition of the recursive splicing site, presumably through the process of exon definition. This is a complex mechanism that defines splice sites on either side of an exon through recruitment of splicing-promoting

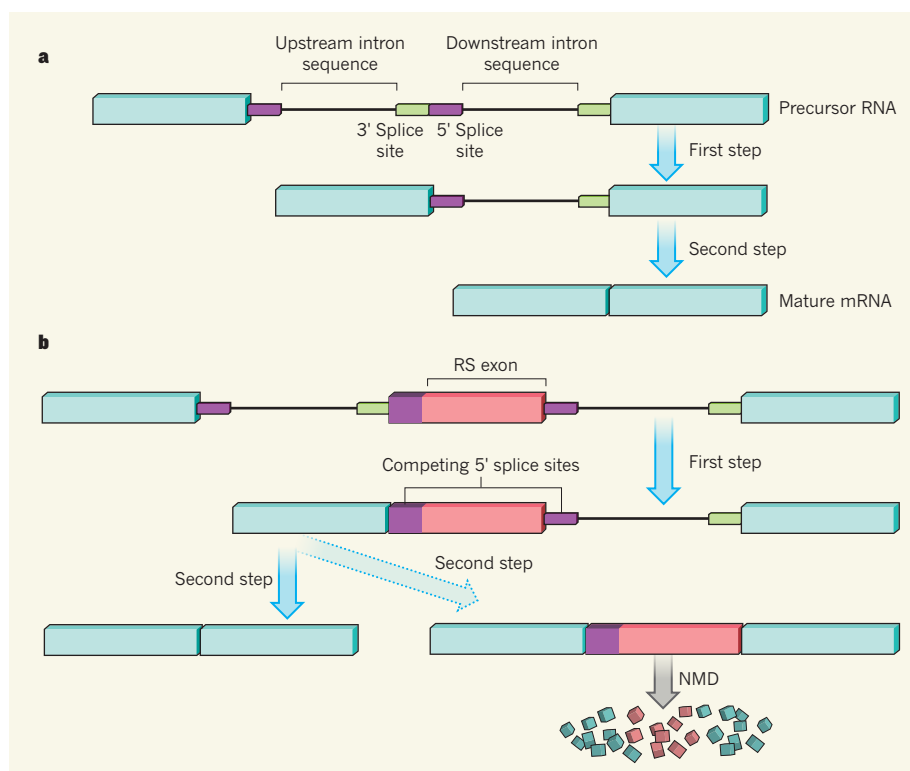


Figure 1 | Mechanisms of recursive splicing. **a**, In recursive splicing, long intron sequences of precursor RNA are removed in a stepwise process mediated by juxtaposed internal 3' and 5' splice sites. In the first step, the 3' splice site is used to remove the upstream intronic sequences. The second step uses the 5' splice site to remove the downstream intron sequences, forming a mature messenger RNA. Duff *et al.*⁵ report that this recursive splicing process occurs in the fruit fly *Drosophila melanogaster* much more commonly than was previously thought. **b**, Sibley *et al.*⁶ find that some recursively spliced messenger RNAs — including all those known in humans — contain a recursive splicing (RS) exon. The RS exon can be either completely removed or retained in the mature mRNA, depending on which of two competing 5' splice sites is used in the second step. Most mRNAs that harbour RS exons are degraded by nonsense-mediated RNA decay (NMD).

proteins⁷. Second, it provides opportunities for quality control: RS exons are almost always spliced out of normal mRNAs, but the authors found that they are usually retained when the upstream exon is generated from an aberrant promoter sequence or from a potentially faulty splicing event. RS-exon inclusion is favoured in these instances because its 5' splice site drives splicing more effectively than the 5' splice site required to remove the RS exon.

RS-exon retention often leads to death of the mRNA, because RS exons typically contain in-frame premature-termination codons — sequences that cause the mRNA to be degraded by the nonsense-mediated RNA decay (NMD) pathway⁸ (Fig. 1b). This is physiologically relevant because most RS-exon-containing mRNAs are probably 'garbage' transcripts. But a subset of these mRNAs may be functional; their formation might be induced when NMD is repressed, such as during particular stages of development and in response to stress⁸.

Why do humans and *Drosophila* seem to use different mechanisms to splice out recursive exons? Species-specific splicing factors may be one explanation. Alternatively, differential RS-exon usage might result from known differences in how these two species define splice

sites⁷. It could also be that the differences in these two species seem greater than is actually the case — for example, RS exons might participate in an intermediate step of *Drosophila* recursive splicing, being included in mature RNAs so infrequently that they are usually undetectable.

It was previously proposed that recursive splicing might increase the fidelity of splicing^{1–3}. Sibley *et al.* examined this possibility using antisense oligonucleotide molecules to block recursive splice sites. They found that this had no obvious effect on the recursive splicing of two human genes, and only modestly inhibited recursive splicing of a zebrafish gene. These data suggest that recursive splicing is not required for the efficiency or accuracy of long-intron splicing. It is possible, however, that this experiment did not reveal a crucial role of recursive splicing because blockade of the natural recursive splice site led to the use of other recursive splice sites that are not normally used.

Duff *et al.* performed extensive genome-wide analyses of *Drosophila* (35 dissected tissues, 24 cell lines and 30 developmental stages) and found that recursive splicing occurs in about 6% of long introns in all tissues tested. By contrast, recursive splicing

may exhibit some tissue specificity in humans. Sibley *et al.* found that genes with long introns tend to be expressed in the human nervous system, and they identified recursively spliced RNAs expressed in the human brain⁶. Duff *et al.* detected some selectivity for recursive splicing in the brain in a screen of 20 human tissues (including fetal brain and adult cerebellum), but this may partly reflect the difficulty of detecting recursively spliced RNAs in tissues that express such RNAs at low levels. It will be important to determine whether this specificity, if real, results from the tendency of recursively spliced genes to be expressed in the brain, or whether cells in the nervous system have factors that promote recursive splicing.

Many genes that have long introns, including those that undergo recursive splicing, are linked to neurological diseases and to autism^{9–11}. Whether these conditions are sometimes triggered by errors in the multi-step recursive RNA-splicing process will be an exciting avenue for future studies. ■

Heidi Cook-Andersen and Miles

F. Wilkinson are in the Department of Reproductive Medicine, University of California, San Diego, La Jolla, California, 92093, USA. M.F.W. is also at the Institute of Genomic Medicine, University of California, San Diego.

e-mail: mfwilkinson@ucsd.edu

1. Hatton, A. R., Subramaniam, V. & Lopez, A. J. *Mol. Cell* **2**, 787–796 (1998).
2. Burnette, J. M., Miyamoto-Sato, E., Schaub, M. A., Conklin, J. & Lopez, A. J. *Genetics* **170**, 661–674 (2005).
3. Grellscheid, S.-N. & Smith, C. W. J. *Mol. Cell. Biol.* **26**, 2237–2246 (2006).
4. Shepard, S., McCreary, M. & Fedorov, A. *PLoS ONE* **4**, e7853 (2009).
5. Duff, M. O. *et al.* *Nature* **521**, 376–379 (2015).
6. Sibley, C. R. *et al.* *Nature* **521**, 371–375 (2015).
7. Berget, S. M. *J. Biol. Chem.* **268**, 2411–2414 (1993).
8. Karam, R., Wengrod, J., Gardner, L. B. & Wilkinson, M. F. *Biochim Biophys Acta* **1829**, 624–633 (2013).
9. Lagier-Tourenne, C. *et al.* *Nature Neurosci.* **15**, 1488–1497 (2012).
10. Polymenidou, M. *et al.* *Nature Neurosci.* **14**, 459–468 (2011).
11. King, I. F. *et al.* *Nature* **501**, 58–62 (2013).

This article was published online on 13 May 2015.

CLARIFICATION

The News & Views article 'Quantum physics: Two-atom bunching' by Lindsay J. LeBlanc (*Nature* **520**, 36–37; 2015) described a paper reporting a type of two-particle quantum interference called the Hong-Ou-Mandel effect using helium-4 atoms, but did not make clear that similar two-particle quantum interference had previously been reported using rubidium-87 atoms (A. M. Kaufman *et al.* *Science* **345**, 306–309; 2014).

The crystallography of correlated disorder

David A. Keen¹ & Andrew L. Goodwin²

Classical crystallography can determine structures as complicated as multi-component ribosomal assemblies with atomic resolution, but is inadequate for disordered systems—even those as simple as water ice—that occupy the complex middle ground between liquid-like randomness and crystalline periodic order. Correlated disorder nevertheless has clear crystallographic signatures that map to the type of disorder, irrespective of the underlying physical or chemical interactions and material involved. This mapping hints at a common language for disordered states that will help us to understand, control and exploit the disorder responsible for many interesting physical properties.

It is remarkable that, for all its indisputable successes, the language of classical crystallography still cannot properly describe the structure of water ice¹. This inadequacy is clearly not a consequence of an especially complicated molecular structure or an enormous unit cell. Nor is it because water cannot be crystallized: snowflakes are as commonly recognized to be crystals as are rock salt and gemstones. Rather, it is because this chemically simple system challenges the very axiom on which classical crystallography rests: namely, the existence of translational periodicity. Though far from random, the orientations of water molecules in ice are not periodic. There is no space group to describe the local correlations that do persist; the very language of classical crystallography fails. Even the higher-dimensional abstractions that describe quasicrystals and incommensurate phases cannot help. Yet the relative orientations of neighbouring water molecules still shape the physical properties of ice because they govern its hydrogen-bonding network, its distribution of electric dipoles, and its anomalous configurational entropy—all of which would differ fundamentally were its structure entirely periodic.

Ice is by no means the only exception to the crystallographic rulebook, and many before us have already commented on the inadequacy of classical crystallography to describe important families of materials^{2,3}. Indeed, one of our aims here is to highlight not only the progress that has been made in probing and understanding correlated disorder, but also the diversity of scientific domains in which it is assuming increasing importance. This diversity is reflected even in the dense packings of simple polyhedra, among which correlated disordered states now appear to be more common than crystalline phases⁴. Yet we are somewhat biased against the study of disordered states because our analytical techniques are so predisposed to order. In this sense the timing of our review is by no means accidental: the convergence of modern developments in crystallographic methods—experimental, computational and algorithmic—now provides the necessary insight into correlated disorder, so that long-standing problems in the field can at last be addressed. Moreover, just as classical crystallographic techniques can be applied across the sciences—to proteins and minerals and magnets alike—and the data understood and interpreted within one theoretical framework, so too do the diffraction patterns of materials with different types of correlated disorder hint at a common descriptive language for these states.

What emerges is that correlated disorder seems to arise in one of two scenarios. There may be an incompatibility between the interactions that drive order and the geometry of the lattice on which that order must evolve; this is the ‘geometric frustration’ well known in the field of frustrated magnetism. Alternatively, the dominant interactions in a system might be satisfied in sufficiently many ways that they do not encode a unique ordering pattern, as illustrated by the square ice model of Box 1 Figure. These two ordering problems—one of frustrated overconstraint and the other of configurational underconstraint—are related in that they both give rise to a degenerate manifold of ground states. This degeneracy results in a set of characteristic physical properties that emerge irrespective of the physical origin of the particular interactions involved: extreme susceptibilities, low-energy excitations, liquid-like behaviour and collective or emergent states. A second consequence is the distinction between local and average symmetry that is so often implicated in the unusual physical properties of disordered crystals (see also Box 1).

We begin our Review by explaining the experimental signature of correlated disorder (that is, diffuse scattering) and how it is measured and interpreted. Drawing on a wide range of examples, we establish qualitative mappings of problems of correlated disorder across ostensibly disparate fields. We show that such mappings concern not only the microscopic nature of correlated disorder but also the form of the diffraction patterns observed in crystallographic experiments. In this way, many of the crystallographic tools conventionally used to study periodic order can help us with the arguably more difficult problem of understanding disorder. Before concluding with a discussion of the opportunities and challenges for researchers in the field, we highlight how that central mantra of structural science—namely, that structure determines function—applies as much to correlated disordered states of matter as it does to crystals.

Scattering from disordered crystals

Crystal structures are determined by interpreting diffraction patterns obtained by scattering X-rays, neutrons or electrons from crystalline samples. The diffraction pattern of an ideal crystal and even of a quasicrystal and of an incommensurately modulated crystal is a set of discrete sharp Bragg reflections, whereas for the most disordered condensed phases (liquids or glasses) the scattering consists of smoothly continuous broad rings with no Bragg peaks at all. Crystal systems with

¹ISIS Facility, Rutherford Appleton Laboratory, Harwell Oxford, Didcot, Oxfordshire OX11 0QX, UK. ²Department of Chemistry, University of Oxford, Inorganic Chemistry Laboratory, South Parks Road, Oxford OX1 3QR, UK.

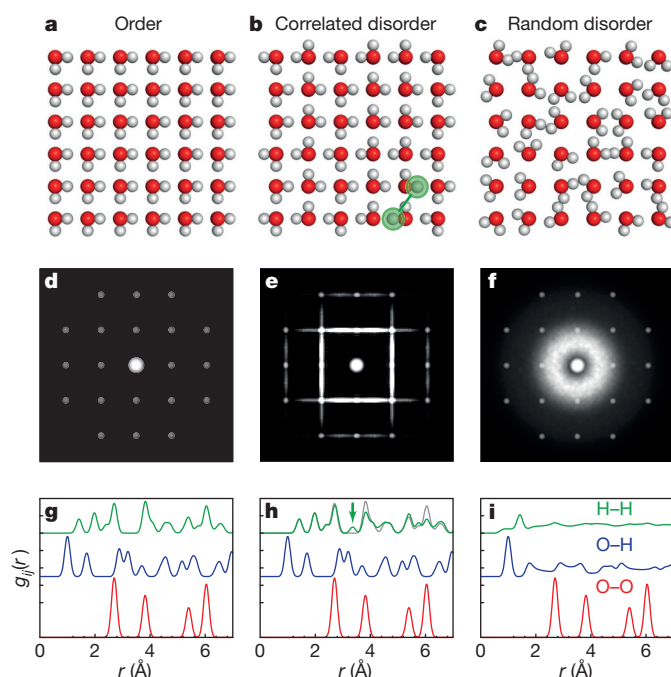
BOX 1

Correlated disorder in square ice

The figure shows how structures with correlated disorder (**b, e, h**) fall in between those with complete order (**a, d, g**) and those that are disordered randomly (**c, f, i**). The differences between the structures of the systems (**a–c**) are evident from the ‘diffuse’ scattering in their diffraction patterns (**d–f**) and from their PDFs (**g–i**). The example used is that of square ice, a simplified two-dimensional representation of the three-dimensional structure of water ice.

Interestingly, this hypothetical square ice structure is similar to a phase very recently observed in thin layers of water trapped between graphene sheets⁷¹. Each oxygen atom (red) on a square grid is surrounded by four hydrogen atoms (white); two of these are covalently bonded and two are hydrogen-bonded to the oxygen atom. This simple rule, developed by Bernal and Fowler⁷² and Pauling⁴⁰, might be expected to lead to an ordered structure (**a**). But the short-range nature of the rule means that structural variants as in **b**, with correlated disorder and residual low-temperature entropy⁷³, are more likely. The correlation in **b** involves all O–H bonds pointing along a common direction for each individual row and column; the structure is disordered because these directions are uncorrelated across rows or columns. The average crystallographic unit cell for square ice has two 50%-occupied hydrogen sites between each pair of oxygen sites. Since the repeating lattice is the same for structures **a–c**, their Bragg scattering, shown as grey dots in the X-ray diffraction plots in **d–f**, is essentially indistinguishable. It is therefore the broad or ‘diffuse’ scattering in plots **d–f**, and the pair distribution functions (plots **g–i**) that are sensitive to the differences between these structures.

No diffuse scattering is observed for ordered structure **a** (plot **d**), whereas horizontal and vertical lines of diffuse scattering are produced by structure **b** from the locally correlated hydrogen atom positions (plot **e**). Structure **b** should be contrasted with a random arrangement of water molecule orientations on a square lattice (**c**), whose average unit cell would have a ring of hydrogen density around each oxygen site; this gives a broad ring of diffuse scattering centred at the origin (as in **f**), more akin to the familiar solvent ring seen in diffraction from proteins. Indeed, a diffuse ring like this would be the only feature in a diffraction pattern from two-dimensional ‘liquid’ water, because the lattice that gives rise to Bragg peaks would have melted. The equivalent partial PDFs are shown in **g–i** for **a–c**, respectively. These functions represent the probability of finding an atom of one type a certain distance away from an atom of another type and local deviations from the ideal lattice are clearly seen. For example, whereas all structures produce identical $g_{OO}(r)$ (red lines) reflecting the regular arrangement of oxygen atoms, subtle additional H–H peaks in **h** but not in **g** (green lines; the $g_{HH}(r)$ from **g** is reproduced in **h** as a thin grey line) are characteristic of specific correlations that exist in structure **b** but not in structure **a**. One such H–H correlation (green arrow in **h**) is highlighted in green in **b**. (The PDFs were calculated with a representative ‘experimental’ Gaussian broadening. Diffuse scattering was calculated using Discus¹⁹.)



correlated disorder produce diffraction patterns that contain discrete Bragg reflections as well as continuous scattering. It is this structured continuous (or ‘diffuse’) scattering that reflects the correlations present in the disordered component, yet classical crystallography gives no recipe for its analysis. If, as is often the case, this component is simply ignored and conventional approaches are used to measure and interpret the Bragg peak intensities, then the structural model obtained represents a configurational average over all possible disordered states and any information concerning correlation is lost (see also discussion in Box 1).

Experimentally, diffuse scattering is collected using methods that survey large swathes of reciprocal space. In the early days, typified by the extensive investigations of Lonsdale^{5,6}, polychromatic (‘Laue’) or monochromatic X-ray beams, single crystals and photographic film were employed in long-exposure experiments to reveal the diffuse features that are typically two to four orders of magnitude weaker than Bragg reflections. Nowadays large-area detectors have replaced film, and X-ray synchrotron instrumentation can collect three-dimensional volumes of high-quality reciprocal-space data extremely rapidly^{7,8}. Similar neutron and electron reciprocal-space survey data can be measured using instruments such as the SXD instrument at ISIS^{9,10} and transmission electron microscopes¹¹, respectively.

A measurement from a crystalline powder or large collection of nanocrystals averages the scattering volume into a one-dimensional

diffraction pattern. Such ‘total scattering’ measurements are the powder equivalent of the surveying methods described above and have become increasingly popular in the past 25 years. They yield absolutely normalized diffraction patterns that contain all Bragg and diffuse scattering—the ‘total scattering structure factor’. The Fourier transform of this function is the pair distribution function (PDF): a weighted sum of partial pair distribution functions, each of which describes the probability of finding atoms of one type a certain distance away from atoms of another type^{12,13}. This very intuitive function (see Box 1) is important for understanding disordered crystal structures since it is one of the few measurable functions that directly accesses the key distances that are longer than those between bonded atoms (which fall within the domain of extended X-ray absorption fine structure and nuclear magnetic resonance measurements) and shorter than those which begin to reflect longer-range periodicity (tractable through analysis of Bragg peak intensities alone). Moreover, measurement and interpretation of a PDF need not draw on the assumptions of complete periodic order inherent to classical crystallography.

A relatively large number of instruments specifically designed for total scattering measurements have been commissioned at neutron sources, X-ray synchrotrons and by commercial companies making diffractometers for laboratory use. This is due to an increased awareness of the scientific merits of a total scattering experiment by crystallographers who

are starting to tackle structural problems that involve correlated disorder and, as a result, have demanded better instrumentation and greater access. Neutron diffractometers at time-of-flight sources produce the highest quality PDFs because they reach the highest possible momentum transfers Q : with the real-space resolution Δr of the PDF proportional to $1/Q_{\max}$ and using $Q_{\max} \approx 50 \text{ \AA}^{-1}$, data reaching a real-space resolution of $\Delta r \approx 0.1 \text{ \AA}$ are routinely measurable. Moreover, these data can be placed on an absolute scale, which allows direct comparison with PDFs determined from candidate structural models. PDF methods have now become part of the toolkit not only for neutron diffraction but also for mainstream crystallography, thanks to the high-energy X-rays available at X-ray synchrotrons that permit routine and rapid collection of high-quality X-ray total scattering data^{14,15}. The latter are limited in comparison to neutron data by the lower accessible values of $Q_{\max} \approx 35 \text{ \AA}^{-1}$ and the much reduced signal at high Q that results from the X-ray form factors. However, as in classical crystallographic studies, it is now increasingly common for neutron and X-ray total scattering data to be analysed in tandem. This is beneficial because different weightings lead in many cases (such as metal organic frameworks¹⁶) to different atom pairs dominating the neutron and X-ray PDFs.

Improved resolution and accuracy, and availability of X-ray and neutron data all help in the interpretation of the 3–10 Å region of the PDF that plays a key part in correlated disordered systems. However, perhaps most important is the development of modelling methods to interpret total scattering patterns.

Methods of analysing diffuse scattering

Interpretation of the diffuse contribution to diffraction patterns of disordered materials has typically been carried out on a case by case basis: experimenters observe diffuse scattering features and develop tools, which are often sample specific, for analysing their measurements with varying degrees of sophistication. Established protocols typically only consider small local defects or modulated deviations from the average periodicity, so such analyses are highly intuitive as there are no generic analytical methods for characterizing correlated disordered states. This said, one approach that has come to prominence recently is the ‘real-space Rietveld’ method implemented within the PDFGui computer program¹⁷ that allows refinement of average-structure-like parameters (such as the unit cell and atom coordinates) against the PDF. Though the resulting structure is based on a periodic unit cell (this is the ‘small box’ of small-box modelling), the parameters are influenced by the local correlations captured in the low- r part of the PDF.

Many methods are limited in that their parameterized structural disorder may or may not be compatible with a physical arrangement of atoms. Analyses that construct or refine a ‘big-box’ model—typically a large supercell of the underlying crystal unit cell—are better able to generate a holistic picture of complex disordered states. These methods have become tractable with increases in computer power and the development of effective strategies for calculating single-crystal diffuse scattering patterns^{18,19}. Monte Carlo calculations can now be tested with^{20,21}, and interaction potential parameters refined against²², experimental diffuse scattering data. Diffuse scattering patterns can also be calculated for configurations generated during molecular dynamics simulations and subsequently compared against experimental data, an approach used in studies of silver-ion conductors²³ and organic semiconductors²⁴. Diffuse scattering calculated²⁵ from *ab initio* phonon energies and eigenvectors has also been compared to experimental data²⁶, and mean-field formulations have been used for magnetic diffuse scattering calculations²⁷.

The reverse Monte Carlo method^{28,29} refines ‘big-box’ models against experimental data by minimizing, using Monte Carlo protocols, the difference between functions calculated from the box of atoms and those determined experimentally. The appeal of the method lies with its generality and flexibility: no symmetry is enforced apart from periodic boundary conditions and the overall shape of the model, and constraints or restraints are only built into the method as needed. Also, any function

that might be calculated from an atomistic model can be incorporated into the calculation of the agreement function. This is especially useful for disordered materials because both the Bragg profile (long-range periodicity) and total scattering functions (short-range disorder) can readily be calculated and used. The resulting refined ‘big-box’ model represents a snapshot of the disordered crystal structure, containing local correlations consistent with the measured total scattering while also replicating the average structure as revealed by the Bragg profile. It is a snapshot of the structure because the interaction between the X-ray (or neutron) and the sample is very fast on the timescale of atomic vibrations and because the total scattering function is predicated on a measurement that integrates over all energy changes in the material. This means that all disordering processes (whether dynamic or static) are captured within the reverse Monte Carlo refined model.

Mapping of apparently unrelated problems

If the presence of correlated disorder gives rise to structured diffuse scattering and/or unexpected pair correlations in the PDF, then it makes sense to investigate the relationship in reverse. That is, to what extent might diffuse scattering patterns or variations in the PDF be diagnostic of specific forms of correlated disorder?

Frustrated overconstraint

Our attempt to answer this question begins by considering the particular type of disorder associated with geometric frustration, typified by the problem of the Ising triangular antiferromagnet. A triangular array of Ising ‘spins’ (which can assume one of only two possible states, ‘up’ and ‘down’) is frustrated if energy is minimized by neighbouring spins adopting opposite states: each pair of opposing neighbours also share a third common neighbour, which cannot oppose both spins in the pair at once (Fig. 1a). The energy minimum for this system corresponds to a compromise situation in which each spin opposes, on average, only four of its six nearest neighbours and is forced to point in the same direction as the other two (hence ‘frustration’). Figure 1b sketches one representative solution, but there are so many equivalent states that the system is characterized by a finite residual entropy at 0 K (ref. 30).

Despite the absence of long-range order amongst this family of ground states, the spin arrangements in a triangular Ising antiferromagnet remain correlated and are not random. Indeed, correlations with certain periodicities are very much stronger than others, resulting in a diffraction pattern that contains stronger scattering at points in reciprocal space that reflect these favoured periodicities. In this case, the scattering is particularly structured because it traces the Brillouin zone boundary of the underlying triangular lattice, reflecting the basic alternation in Ising state from one lattice site to the next. The absence of Bragg reflections means that no structure solution is possible from this scattering pattern using conventional techniques and there is no ‘unit cell’ that can describe the spin arrangement. Instead, a direct Fourier transform of the scattering function gives what is perhaps the closest equivalent: a two-dimensional spin correlation function representing the ‘average local structure’ of the disordered state³¹.

The Ising model can be equally meaningfully applied across a range of systems beyond the field of magnetism. A problem of vacancy ordering, for example, might be described by the distribution of occupied (Ising ‘up’) and vacant (Ising ‘down’) sites, with the corresponding ‘antiferromagnetic’ interaction involving strain-driven anti-clustering of vacancies. Likewise, the arrangement of cations of two different formal charges (for example, $2+ = \text{‘up’}$, $3+ = \text{‘down’}$) might be dictated by the stronger Coulombic repulsion of the more highly charged component. In both cases one anticipates the realization of complex states in which the vacancy or charge localization can be mapped onto the spin states of the triangular Ising antiferromagnet, with the corresponding diffraction patterns—if measured using a probe of relevant sensitivity—reflecting the same structured diffuse scattering pattern associated with that canonical system. In this way, the form of the diffuse scattering pattern becomes a characteristic signature of a particular type of correlated

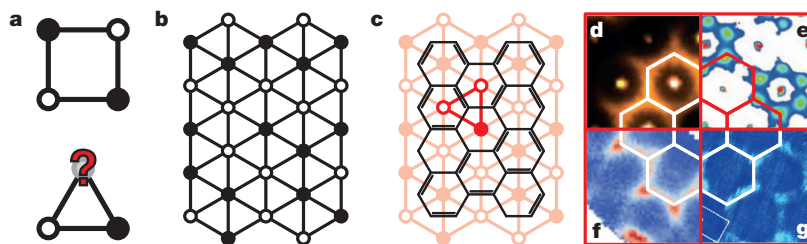


Figure 1 | Correlated disorder in the Ising triangular lattice. **a**, A lattice of sites in one of two states (such as spin up or spin down; here, black or white), adopts a complex structure wherever there is an incompatibility between the interaction of neighbours and the lattice geometry. ‘Antiferromagnetic’ interactions—favouring a reversal in state between neighbours—are easily satisfied on a square lattice (top) but cannot be fully satisfied on a triangular lattice (bottom). **b**, The compromise that emerges for an extended triangular lattice is for each triangle to adopt either a one-up-two-down or a one-down-two-up arrangement; one possible solution is shown here, though many others exist. **c**, There is a one-to-one mapping between these Ising states and the possible arrangements of π bonds in the resonating valence bond model for graphene: the π bonds are simply associated with the ‘ferromagnetic’ pairs on each triangle. **d–g**, Such mappings mean that qualitatively similar scattering functions are observed for structurally complex systems wherever that complexity has the same fundamental geometric origin. The packing arrangement of the Gag protein⁸ (**d**), the orbital orientations in $\text{Ba}_3\text{CuSb}_2\text{O}_9$ (ref. 32) (**e**), the emergent spin structure of β -Mn (ref. 27) (**f**) and the electronic structure of graphene itself³³ (**g**) all map onto the problem of the Ising triangular antiferromagnet. Here, their X-ray (**d**, **e**), polarized neutron (**f**) and angle-resolved photoemission (**g**) scattering patterns all reveal the population of configurational, magnetic or electronic states with periodicities at the Brillouin zone boundary (solid white or red lines).

disorder that in itself might be taken to reflect a conceptual mapping among related problems of geometric frustration.

Such mappings are borne out in practice. In its crystal structure, the amino-terminal fragment of the Gag protein from Feline Foamy Virus packs on a triangular lattice with neighbouring fragments oriented in opposite directions in order to maximise packing efficiency⁸. Here the frustration involves fragment orientations, and the mapping onto the Ising problem is evident in the form of the structured diffuse scattering in X-ray diffraction patterns (Fig. 1d). In Co-doped β -Mn, ferromagnetically coupled multi-spin rods behave as collective magnetic entities that are themselves antiferromagnetically coupled. Because the multi-spin rods are arranged on a triangular lattice, the emergent spin structure is frustrated, with only the part of the spin-polarized single-crystal neutron scattering pattern sensitive to the magnetic structure containing diffuse scattering at the zone boundary (Fig. 1f)²⁷. Geometric frustration of Jahn–Teller-distorted Cu^{2+} ions in $\text{Ba}_3\text{CuSb}_2\text{O}_9$ actually maps the problem of orbital order in this system onto the same Ising configurations, giving rise to a qualitatively similar scattering function (Fig. 1e)³². A final example involves the electronic band structure of graphene and its analogues: the well-established correspondence between the resonating valence bond model and the Ising triangular antiferromagnet (Fig. 1c) means that the momentum distribution of bonding states observed using angular-resolved photoemission spectroscopy also reveals the same distinctive pattern in reciprocal space (Fig. 1g)³³.

Despite the obvious similarity in the scattering patterns of these diverse systems, there remain important differences in the scattering detail. On the very simplest level, it is known that the positions of the scattering maxima reflect the nature and strength of next-nearest-neighbour correlations; likewise the widths of the features are sensitive to the concentration of Ising excitations⁸. It is in this sense that the structure refinement tools outlined in ‘Methods of analysing diffuse scattering’ can provide useful physical insight beyond the qualitative mappings of Fig. 1. Some of this sensitivity remains even for polycrystalline samples; the corresponding PDFs (or spin correlation functions for magnetic systems) are sensitive to longer-range interactions and may provide a direct measurement of correlation length³⁴.

Geometric frustration can and does emerge in many different systems with varying types of degrees of freedom and interactions well beyond those of the Ising triangular antiferromagnet. Historically, much of the emphasis has been on frustrated magnetism, because even within this individual field there are so many types of spin system (such as Ising, XY, Potts and Heisenberg), spin interaction (such as ferromagnetic, antiferromagnetic and bilinear–biquadratic), and lattice geometries (such as triangular, kagome, hyperkagome, pyrochlore) that the variety of complex magnetic states accessible in theory is large indeed³⁵. Yet the part

played by geometric frustration in governing a range of other types of ordering phenomena is becoming increasingly clear: examples include orbital order in the colossal magnetoresistance manganites³⁶, orientational correlations in plastic crystals³⁷, and collective transport properties of superionics³⁸ and elements alike³⁹. As for the Ising triangular antiferromagnet, what emerges is that the underlying lattice geometry influences both the particular type of configurational degeneracy and also the basic form of the diffuse scattering observed experimentally.

Configurational underconstraint

This duality of conceptual and empirical mappings is often illustrated using materials based on the pyrochlore lattice of connected tetrahedra. In many ways cubic ice can be considered the parent of this family of materials: the relationship between its structure and that of the pyrochlore lattice can be seen by considering the network formed by the midpoints between connected oxygen atoms (Fig. 2a). The need for water molecules to be arranged so that each oxygen atom is connected to four hydrogen atoms, two of them covalently bonded and the other two hydrogen-bonded, results⁴⁰ in a disordered arrangement of H_2O molecule orientations, with each arrangement uniquely characterized by the choice of one particular edge per pyrochlore tetrahedron (shown in bold in Fig. 2a). It is a direct consequence of the ice rules that none of these edges meet; there is a one-to-one correspondence between unconnected edge decorations of the pyrochlore lattice and cubic ice configurations. In a similar way, both the direction of Cd^{2+} ion displacements in cubic $\text{Cd}(\text{CN})_2$ and the arrangement of Ho^{3+} spins in $\text{Ho}_2\text{Ti}_2\text{O}_7$ can be mapped onto these same edge decorations (Fig. 2b and c). These three systems then share the same type of configurational degeneracy, and the terms ‘charge ice’ and ‘spin ice’ are used to describe the families of materials of which $\text{Cd}(\text{CN})_2$ (ref. 41) and $\text{Ho}_2\text{Ti}_2\text{O}_7$ (ref. 42) are examples. Once again, these mappings amongst equivalent disordered states are reflected in the experimental diffuse scattering patterns. Water, charge, and spin ices all give rise to the same characteristic type of scattering pattern (Fig. 2e–g), with higher-order correlations and particulars of the scattering physics affecting the distribution of scattering intensity throughout reciprocal space⁴³. That a qualitatively similar scattering pattern is associated with some superionic conductors (Fig. 2d) suggests that local correlations also have a role in these materials, even if the specific mechanism involved remains controversial⁴⁴.

Symmetry mismatch and local disorder

What is common to all these systems is that there exists a distinction between local symmetry and the (higher) average symmetry imposed by the crystal lattice. Individual water molecules have lower point symmetry than the tetrahedral crystallographic sites on which they sit in solid ice; likewise, the two-up-one-down compromise of the Ising antiferromagnet

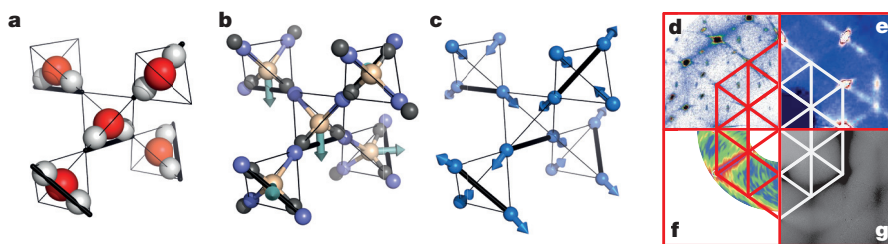


Figure 2 | Ice-like states on the pyrochlore lattice. **a**, The structure of cubic ice is related to that of the pyrochlore lattice (thin black lines). Many different orientations of the water molecules are capable of satisfying the same hydrogen-bonding ‘rules’, in which each configuration can be represented uniquely by decorating the edges associated with the two hydrogen atoms per tetrahedron (thick black lines). The ice rules are encoded in the provision that no two such edges join. **b**, **c**, The arrangement of dipoles in ‘charge-ices’ (**b**) and magnetic

moments in ‘spin-ices’ (**c**) map onto the same edge decorations, linking the structural complexity of these physically disparate systems. **d–g**, Diffraction patterns of the superionic conductor α -Cu_{1.8}Se (ref. 67) (**d**), the negative thermal expansion ‘charge-ice’ Cd(CN)₂ (ref. 41) (**e**), the quantum spin ice candidate Yb₂Ti₂O₇⁶⁸ (**f**) and (water) ice itself⁶⁹ (**g**) all show continuous scattering in related regions of reciprocal space.

breaks the threefold symmetry of the triangular lattice. The impact of this symmetry mismatch is for the average structure—deduced from analysis of Bragg diffraction—to appear to have a higher symmetry than that observed using local spectroscopic probes or expected from crystal chemical considerations. This has led to misconceptions about disordered crystal structures in the past (for example, the apparently linear Si–O–Si bond in β -cristobalite⁴⁵), and the clearest way to address this ambiguity is through direct analysis of the diffuse scattering or through modelling of the total scattering. Big-box modelling is particularly effective because local distortions within individual unit cells can be explored (to agree with PDF data, for example) while still generating an overall structure that replicates the Bragg diffraction intensities.

Such an approach has been used to address high-temperature behaviour in ferroelectric BaTiO₃: even within the cubic (paraelectric) phase the Ti atoms are displaced from the centre of the TiO₆ octahedron in one of the eight $\langle 111 \rangle$ directions, mimicking the distorted arrangement of the ordered low-temperature rhombohedral phase⁴⁶. These eight Ti positions average to the central octahedral B-site in the ideal perovskite structure (Fig. 3a)⁴⁷. A similar conclusion was reached in studies of the thermoelectric properties of PbTe (which has the same structure as rock salt) in terms of large-amplitude displacements of Pb atoms along $\langle 100 \rangle$ -type directions at high temperature⁴⁸. Likewise, the high ionic conductivity of δ -Bi₂O₃ was shown to depend on local relaxation of the Bi-ion coordination geometry towards that adopted in the low-temperature β -phase; these correlated distortions promote vacancy migration (Fig. 3b)⁴⁹. The Bi atoms in the relaxor ferroelectric NBT (Na_{0.5}Bi_{0.5}TiO₃) also assume positions of lower local symmetry than that of the average lattice in the rhombohedral phase⁵⁰. These are all examples where second-order Jahn–Teller distortions of the Bi³⁺, Pb²⁺ and Ti⁴⁺ coordination environments are responsible for lowering the local symmetry. Analogous behaviour is observed for molecular systems where a phase transition can only reduce the overall structural distortion and raise the average symmetry through a superposition of distinct molecular orientations. An example of this is seen in Fig. 3c where rotational disorder in the imidazolium cation (C₃H₅N₂)⁺ above the ferroelectric phase transition leads to an average hexagonal molecular shape that is chemically nonsensical⁵¹.

For most materials correlated disorder persists only in a high-temperature state, with order emerging on cooling. But in some systems the disordered state is trapped to low temperatures. One example is K_{1-x}(NH₄)_xI (with $x \approx 0.5$), where the tetrahedral geometry of the ammonium cation is incompatible with the octahedral symmetry of its crystallographic site in the rock salt structure⁵². A second example is solid C₆₀, where the combination of icosahedral molecular symmetry and trigonal point symmetry at the crystallographic site frustrates order and gives rise to glassy dynamics at low temperatures⁵³. A similar mismatch between symmetry at the molecular and crystal lattice level

also exists in far larger structures, and is even exploited in the mechanical release of phage DNA from viruses⁵⁴.

Disorder–property relationships

Perhaps counterintuitively, correlated disorder may actually be an essential ingredient for functional material properties. There will be even more cases where disorder—though not by itself the microscopic driving force—is intimately associated with a particular functionality. Any switchable ferroic state, for example, emerges from a disordered parent phase where the correlations that are present describe the ferroic property of interest. Relaxor ferroelectrics are an extreme example of this relationship, where correlations are so strong that they stabilize polar nanoregions, which in turn drive the attractive dielectric properties for which relaxors are favoured⁵⁵. Here it is dipolar disorder that results in function, but there are strong analogies too to the balance of orbital, electronic, and magnetic disorder implicated in the colossal magnetoresistance of La_xCa_{1-x}MnO₃, for example⁵⁶. Likewise, the proximity of correlated paramagnetic states to the superconducting transitions of most high-temperature superconductors has been noted many times previously (for example, see ref. 57). As these examples illustrate, there is at least an empirical correspondence between correlated disorder and advanced function that is increasingly obvious even if not yet well understood.

Correlated disorder is often implicated in cooperative phenomena. One example is solid-phase ion conduction: superionics are effectively porous to a particular type of ion precisely because there exists a low-barrier mechanism for collective displacements. Another example is the emergence of quasi-particles such as skyrmions in chiral ferromagnets⁵⁸ and magnetic ‘monopoles’ in the pyrochlore spin ices⁵⁹, and the potential application of these phenomena in data storage and spintronic devices would constitute putting correlated disorder to practical use. As a final point, we note that the configurational entropy associated with disordered states has its own set of thermodynamic and lattice dynamical consequences that affect material properties. Not only would ice melt at a different temperature were it not for proton disorder, but the influence of disorder on phonons is exploited in optimizing thermal conductivity of thermoelectrics⁶⁰.

Concluding remarks

As we have shown here, simple local rules or distortions can give rise to surprisingly complex disordered states of matter in a wide range of material systems. In many cases the presence of disorder—whether chemical, electronic, magnetic or geometric—and the nature of the correlations that persist within the disordered phase affect the physical and chemical properties of the system in question. Just as classical crystallography has helped us to develop an understanding of the structures of ordered crystalline materials, so modern crystallography and its ever-improving methods are rapidly improving our ability to characterize

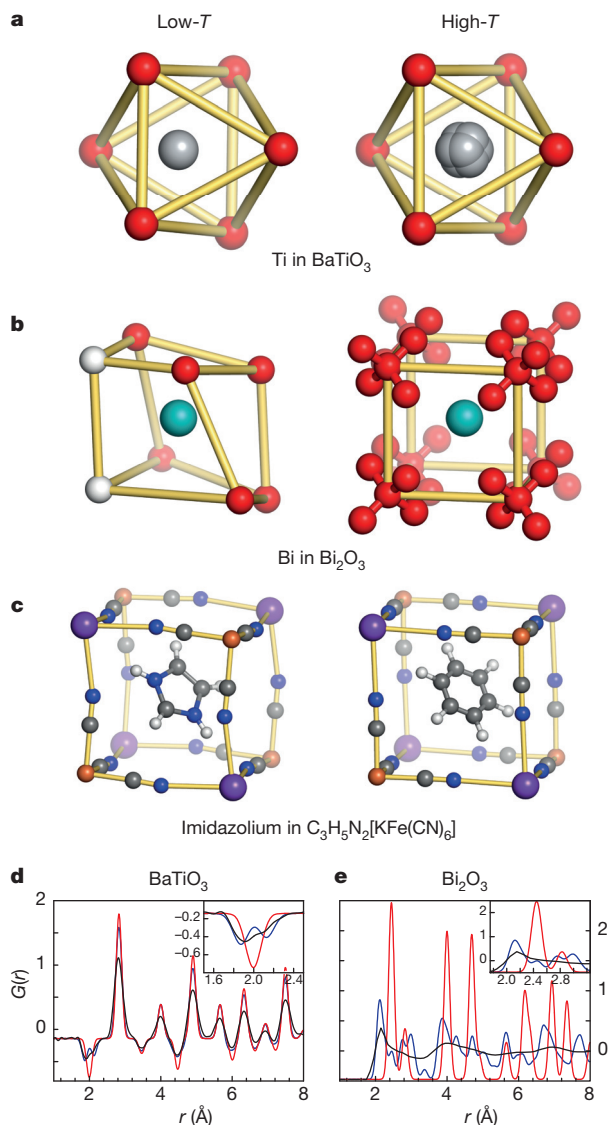


Figure 3 | Local symmetry mismatch. a–c, Examples of phase transitions where mismatch is resolved at low temperature T through lowering of the lattice symmetry and high-temperature average structures rely on descriptions involving partially occupied sites (see main text for details). a, Ti in BaTiO₃; b, Bi in Bi₂O₃; and c, imidazolium in (C₃H₅N₂)₂[KFe(CN)₆]. The experimental pair distribution functions $G(r)$ (black lines) from a and b at high temperature are compared with the $G(r)$ values calculated using PDFGui¹⁷ from the average structures of the phases at low (blue lines) and high (red lines) temperatures in d and e, respectively. In each case the low- r portion of the measured high-temperature phase $G(r)$ more closely resembles that calculated from the average structure of the low-temperature phase (see insets to d and e). Reference structures are: a, average structures⁷⁰ for the low-temperature $R3m$ phase and the high-temperature cubic $Pm\bar{3}m$ phase; b, average structures at 297 K ($P4_21c$, β -phase) and 1,033 K ($Fm\bar{3}m$, δ -phase)⁴⁹; c, for the order–disorder transition from (C₃H₅N₂)₂[KFe(CN)₆] at 83 K ($C2/c$) and 293 K ($R3m$)⁵¹; d, PDF data (J. A. Hriljac and D.A.K., unpublished work); and e, PDF data⁴⁹.

and interpret correlated disorder. We take encouragement from the observation that certain types of correlated disorder recur in completely different fields. This recurrence of specific forms of disorder hints at the possibility of a universal language for describing correlated disordered states, much like the space groups of classical crystallography.

Beyond even the problem of correlated disorder lie exotic states of matter with varying degrees of order and disorder over vastly different (and often multiple) length scales. Meeting the challenge of structurally characterizing such phases will require further development of the ‘small angle’ scattering techniques that probe larger-scale structures.

Ever larger structural models will be needed to aid the interpretation of data from instruments such as NIMROD at ISIS⁶¹ that can measure total scattering to much smaller values of Q . Furthermore, as correlated disorder is investigated in ever more complicated systems, there will be an increasing demand for analysis based on effective interfacing of different techniques: the combined analysis of extended X-ray absorption fine structure, total scattering and single-crystal diffuse scattering will reveal more than each type of data could reveal individually. This convergence of different methods of analysis is beginning to happen⁶², and comparative PDF and single-crystal diffuse scattering studies of molecular systems^{63,64} have highlighted the value of programs^{19,29} able to incorporate both data types⁶⁵. The possibility of extracting the three-dimensional PDF from single-crystal measurements is a particularly exciting development in the field⁶⁶, because it removes the limitation of orientational averaging inherent to all powder methods.

Ultimately, of course, the goal will be to control and exploit correlated disorder. This reverses the paradigm of seeking to understand the disorder responsible for interesting physical properties to one of intentionally employing it as a design element in its own right, in order to engineer materials with novel functionalities. But the crucial first step towards that goal is developing the ability to fully characterize correlated disorder, and we hope that our review has shown how such studies might be initiated and will encourage others to join in with this important branch of modern crystallography.

Received 31 July 2014; accepted 13 March 2015.

- Kuhs, W. F. & Lehmann, M. S. The structure of ice-Ih. *Water Sci. Rev.* **2**, 1–66 (1986).
- Cartwright, J. H. E. & Mackay, A. L. Beyond crystals: the dialectic of materials and information. *Phil. Trans. R. Soc. Lond. A* **370**, 2807–2822 (2012).
- Bernal, J. D. & Carlisle, C. H. The range of generalised crystallography. *Sov. Phys. Crystallogr.* **13**, 811–831 (1969).
- Damasceno, P. F., Engel, M. & Glotzer, S. C. Predictive self-assembly of polyhedra into complex structures. *Science* **337**, 453–457 (2012).
- Lonsdale, K. & Smith, H. An experimental study of diffuse X-ray reflexion by single crystals. *Proc. R. Soc. A* **179**, 8–50 (1941).
- Lonsdale, K. X-ray study of crystal dynamics: an historical and critical survey of experiment and theory. *Proc. Phys. Soc.* **54**, 314–353 (1942).
- Maisel, S. B., Schindzielorz, N., Müller, S., Reichert, H. & Bosak, A. An accidental visualization of the Brillouin zone in an Ni–W alloy via diffuse scattering. *J. Appl. Cryst.* **46**, 1211–1215 (2013).
- Welberry, T. R., Heerdegen, A. P., Goldstone, D. C. & Taylor, I. A. Diffuse scattering resulting from macromolecular frustration. *Acta Crystallogr. B* **67**, 516–524 (2011).
- This elegantly describes how a complicated diffuse scattering pattern might be explained in terms of simple disordering rules.
- Welberry, T. R. *et al.* Diffuse neutron scattering in benzil, C₁₄D₁₀O₂, using the time-of-flight Laue technique. *J. Appl. Cryst.* **36**, 1440–1447 (2003).
- Keen, D. A., Gutmann, M. J. & Wilson, C. C. SXD—the single crystal diffractometer at the ISIS Spallation Neutron Source. *J. Appl. Cryst.* **39**, 714–722 (2006).
- Withers, R. L. Disorder, structured diffuse scattering and the transmission electron microscope. *Z. Kristallogr.* **220**, 1027–1034 (2005).
- Soper, A. K. & Barney, E. R. Extracting the pair distribution function (PDF) from white beam X-ray total scattering data. *J. Appl. Cryst.* **44**, 714–726 (2011).
- Keen, D. A. A comparison of various commonly used correlation functions for describing total scattering. *J. Appl. Cryst.* **34**, 172–177 (2001).
- Chupas, P. J. *et al.* Rapid-acquisition pair distribution function (RA-PDF) analysis. *J. Appl. Cryst.* **36**, 1342–1347 (2003).
- Skinner, L. B., Benmore, C. J. & Parise, J. B. Area detector corrections for high quality synchrotron X-ray structure factor measurements. *Nucl. Instrum. Meth. A* **662**, 61–70 (2012).
- Bennett, T. D. *et al.* Structure and properties of an amorphous metal–organic framework. *Phys. Rev. Lett.* **104**, 115503 (2010).
- Farrow, C. L. *et al.* PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals. *J. Phys. Condens. Matter* **19**, 335219 (2007).
- Butler, B. D. & Welberry, T. R. Calculation of diffuse scattering from simulated disordered crystals: a comparison with optical transforms. *J. Appl. Cryst.* **25**, 391–399 (1992).
- Neder, R. B. & Proffen, T. *Diffuse Scattering and Defect Structure Simulations: A Cook Book using the Program DISCUS* (Oxford Univ. Press, 2008).
- Goossens, D. J., Heerdegen, A. P., Chan, E. J. & Welberry, T. R. Monte Carlo modeling of diffuse scattering from single crystals: the program ZMC. *Metall. Mater. Trans. A* **42**, 23–31 (2011).
- Welberry, T. R., Chan, E. J., Goossens, D. J. & Heerdegen, A. P. Use of Monte Carlo simulation for the interpretation and analysis of diffuse scattering. *Phase Transit.* **83**, 80–98 (2010).

22. Chan, E. J. *et al.* Single-crystal diffuse scattering studies on polymorphs of molecular crystals. I. The room-temperature polymorphs of the drug benzocaine. *Acta Crystallogr. B* **65**, 382–392 (2009).
23. Welberry, T. R. & Paściak, M. Monte Carlo and molecular dynamics simulation of disorder in the Ag⁺ fast ion conductors pearceite and polybasite. *Metall. Mater. Trans. A* **42**, 6–13 (2011).
24. Eggeman, A. S., Illig, S., Troisi, A., Sirringhaus, H. & Midgley, P. A. Measurement of molecular motion in organic semiconductors by thermal diffuse electron scattering. *Nature Mater.* **12**, 1045–1049 (2013).
25. Xu, R. & Chiang, T. C. Determination of phonon dispersion relations by X-ray thermal diffuse scattering. *Z. Kristallogr.* **220**, 1009–1016 (2005).
26. Gutmann, M. J. *et al.* Room temperature single-crystal diffuse scattering and ab initio lattice dynamics in CaTiSiO₅. *J. Phys. Condens. Matter* **25**, 315402 (2013).
27. Paddison, J. A. M. *et al.* Emergent frustration in Co-doped β-Mn. *Phys. Rev. Lett.* **110**, 267207 (2013).
- Reverse Monte Carlo analysis of an experiment uncovers an unexpected magnetic structure with a simple underlying basis.**
28. McGreevy, R. L. Reverse Monte Carlo modelling. *J. Phys. Condens. Matter* **13**, R877–R913 (2001).
29. Tucker, M. G., Keen, D. A., Dove, M. T., Goodwin, A. L. & Hui, Q. RMCProfile: reverse Monte Carlo for polycrystalline materials. *J. Phys. Condens. Matter* **19**, 335218 (2007).
30. Wannier, G. H. Antiferromagnetism. The triangular Ising net. *Phys. Rev.* **79**, 357–364 (1950).
31. Paddison, J. A. M., Stewart, J. R. & Goodwin, A. L. SPINVERT: a program for refinement of paramagnetic diffuse scattering data. *J. Phys. Condens. Matter* **25**, 454220 (2013).
32. Nakatsuji, S. *et al.* Spin-orbital short-range order on a honeycomb-based lattice. *Science* **336**, 559–563 (2012).
33. Rasche, B. *et al.* Stacked topological insulator built from bismuth-based graphene sheet analogues. *Nature Mater.* **12**, 422–425 (2013).
34. Paddison, J. A. M. *et al.* Spin correlations in Ca₃Co₂O₆: polarized-neutron diffraction and Monte Carlo study. *Phys. Rev. B* **90**, 014411 (2014).
35. Lacroix, C., Mendels, P. & Mila, F. *Introduction to Frustrated Magnetism: Materials, Experiments, Theory* (Springer, 2011).
36. Lynn, J. W. *et al.* Order and dynamics of intrinsic nanoscale inhomogeneities in manganites. *Phys. Rev. B* **76**, 014437 (2007).
37. Shintani, H. & Tanaka, H. Frustration on the way to crystallization in glass. *Nature Phys.* **2**, 200–206 (2006).
38. Wood, B. C. & Marzari, N. Dynamical structure and thermodynamics of the superionic sublattice in α-AgI. *Phys. Rev. Lett.* **97**, 166401 (2006).
39. Ogitsu, T., Schwegler, E. & Galli, G. β-rhombohedral boron: at the crossroads of the chemistry of boron and the physics of frustration. *Chem. Rev.* **113**, 3425–3449 (2013).
40. Pauling, L. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *J. Am. Chem. Soc.* **57**, 2680–2684 (1935).
41. Fairbank, V. E., Thompson, A. L., Cooper, R. I. & Goodwin, A. L. Charge-ice dynamics in the negative thermal expansion material Cd(CN)₂. *Phys. Rev. B* **86**, 104113 (2012).
42. Bramwell, S. T. *et al.* Spin correlations in Ho₂Ti₂O₇: a dipolar spin ice system. *Phys. Rev. Lett.* **87**, 047205 (2001).
- One of the first experimental papers that initiated the studies of spin ices and subsequent searches for magnetic monopoles.**
43. Fennell, T. *et al.* Magnetoelastic excitations in the pyrochlore spin liquid Tb₂Ti₂O₇. *Phys. Rev. Lett.* **112**, 017203 (2014).
44. Madden, P. A., O'Sullivan, K. F. & Chiarotti, G. Ordering of the silver ions in α-AgI: a mechanism for the α-β phase transition. *Phys. Rev. B* **45**, 10206–10212 (1992).
45. Tucker, M. T., Squires, M. P., Dove, M. T. & Keen, D. A. Dynamic structural disorder in cristobalite: neutron total scattering measurement and reverse Monte Carlo modelling. *J. Phys. Condens. Matter* **13**, 403–423 (2001).
46. Levin, I., Krayzman, V. & Woicik, J. C. Local structure in perovskite (Ba,Sr)TiO₃: Reverse Monte Carlo refinements from multiple measurement techniques. *Phys. Rev. B* **89**, 024106 (2014).
47. Comès, R., Lambert, M. & Guinier, A. Désordre linéaire dans les cristaux (cas du silicium, du quartz, et des pérovskites ferroélectriques). *Acta Crystallogr. A* **26**, 244–254 (1970).
48. Božin, E. S. *et al.* Entropically stabilized local dipole formation in lead chalcogenides. *Science* **330**, 1660–1663 (2010).
49. Hull, S. *et al.* Neutron total scattering study of the delta and beta phases of Bi₂O₃. *Dalton Trans.* **2009**, 8737–8745 (2009).
50. Keeble, D. S. *et al.* Bifurcated polarization rotation in bismuth-based piezoelectrics. *Adv. Funct. Mater.* **23**, 185–190 (2013).
- 'Big-box' modelling of total scattering shows how the local polarization changes within a higher-symmetry average lattice.**
51. Zhang, W., Cai, Y., Xiong, R.-G., Yoshikawa, H. & Awaga, K. Exceptional dielectric phase transitions in a perovskite-type cage compound. *Angew. Chem. Int. Ed.* **49**, 6608–6610 (2010).
52. Berret, J.-F., Bostoen, C. & Hennion, B. Phase diagram of the dipolar glass K_{1-x}(NH₄)_xl. *Phys. Rev. B* **46**, 13747–13750 (1992).
53. David, W. I. F., Ibberson, R. M., Dennis, T. J. S., Hare, J. P. & Prassides, K. Structural phase transitions in the fullerene C₆₀. *Europhys. Lett.* **18**, 219–225 (1992).
54. Liu, X. *et al.* Structural changes in a marine podovirus associated with release of its genome into *Prochlorococcus*. *Nature Struct. Mol. Biol.* **17**, 830–836 (2010).
55. Xu, G., Wen, J., Stock, C. & Gehring, P. M. Phase instability induced by polar nanoregions in a relaxor ferroelectric system. *Nature Mater.* **7**, 562–566 (2008).
56. Billinge, S. J. L., Proffen, Th., Petkov, V., Sarrao, J. L. & Kycia, S. Evidence for charge localization in the ferromagnetic phase of La_{1-x}Ca_xMnO₃ from high real-space-resolution x-ray diffraction. *Phys. Rev. B* **62**, 1203–1211 (2000).
57. Bennemann, K. H. & Ketterson, J. B. (eds) *Superconductivity: Conventional and Unconventional Superconductors* Vol. 1 (Springer, 2008).
58. Mühlbauer, S. *et al.* Skyrmion lattice in a chiral magnet. *Science* **323**, 915–919 (2009).
59. Bramwell, S. T. *et al.* Measurement of the charge and current of magnetic monopoles in spin ice. *Nature* **461**, 956–959 (2009).
60. Christensen, M. *et al.* Avoided crossing of rattler modes in thermoelectric materials. *Nature Mater.* **7**, 811–815 (2008).
61. Bowron, D. T. *et al.* NIMROD: The Near and Intermediate Range Order Diffractometer of the ISIS second target station. *Rev. Sci. Instrum.* **81**, 033905 (2010).
62. Cliffe, M. J. *et al.* Correlated defect nanoregions in a metal–organic framework. *Nature Commun.* **5**, 4176 (2014).
63. Rademacher, N., Daemen, L. L., Chronister, E. L. & Proffen, T. Pair distribution function analysis of molecular compounds: significance and modeling approach discussed using the example of *p*-terphenyl. *J. Appl. Cryst.* **45**, 482–488 (2012).
64. Goossens, D. J. & Whitfield, R. E. Distinguishing types of disorder in diffuse scattering: a numerical simulation study. *Metall. Mater. Trans. A* **45**, 152–161 (2014).
65. Krayzman, V. & Levin, I. Reverse Monte Carlo refinements of nanoscale atomic correlations using powder and single-crystal diffraction data. *J. Appl. Cryst.* **45**, 106–112 (2012).
66. Weber, T. & Simonov, A. The three-dimensional pair distribution function analysis of disordered single crystals: basic concepts. *Z. Kristallogr.* **227**, 238–247 (2012).
67. Danilkin, S. A. Diffuse scattering and lattice dynamics of superionic copper chalcogenides. *Solid State Ionics* **180**, 483–487 (2009).
68. Chang, L.-J. *et al.* Higgs transition from a magnetic Coulomb liquid to a ferromagnet in Yb₂Ti₂O₇. *Nature Commun.* **3**, 992 (2012).
69. Wehinger, B. *et al.* Diffuse scattering in Ih ice. *J. Phys. Condens. Matter* **26**, 265401 (2014).
- High-quality X-ray diffuse and inelastic scattering and computer simulations are combined to show how the molecules in water ice disorder.**
70. Kwei, G. H., Lawson, A. C., Billinge, S. J. L. & Cheong, S.-W. Structures of the ferroelectric phases of barium titanate. *J. Phys. Chem.* **97**, 2368–2377 (1993).
71. Algara-Siller, G. *et al.* Square ice in graphene nanocapillaries. *Nature* **519**, 443–445 (2015).
72. Bernal, J. D. & Fowler, R. H. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J. Chem. Phys.* **1**, 515–548 (1933).
73. Lieb, E. H. Residual entropy of square ice. *Phys. Rev.* **162**, 162–172 (1967).

Acknowledgements We are grateful for discussions with J. A. M. Paddison, M. J. Cliffe and A. B. Cairns. A.L.G. acknowledges financial support from the ERC (grant number 279705).

Author Contributions D.A.K. and A.L.G. contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.A.K. (david.keen@sfc.ac.uk) or A.L.G. (andrew.goodwin@chem.ox.ac.uk).

3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya

Sonia Harmand^{1,2,3}, Jason E. Lewis^{1,3,4}, Craig S. Feibel^{3,4,5}, Christopher J. Lepre^{3,5,6}, Sandrine Prat^{3,7}, Arnaud Lenoble^{3,8}, Xavier Boës^{3,7}, Rhonda L. Quinn^{3,5,9}, Michel Brenet^{8,10}, Adrian Arroyo², Nicholas Taylor^{2,3}, Sophie Clément^{3,11}, Guillaume Daver¹², Jean-Philip Brugal^{3,13}, Louise Leakey¹, Richard A. Mortlock⁵, James D. Wright⁵, Sammy Lokorodi³, Christopher Kirwa^{3,14}, Dennis V. Kent^{5,6} & Hélène Roche^{2,3}

Human evolutionary scholars have long supposed that the earliest stone tools were made by the genus *Homo* and that this technological development was directly linked to climate change and the spread of savannah grasslands. New fieldwork in West Turkana, Kenya, has identified evidence of much earlier hominin technological behaviour. We report the discovery of Lomekwi 3, a 3.3-million-year-old archaeological site where *in situ* stone artefacts occur in spatio-temporal association with Pliocene hominin fossils in a wooded palaeoenvironment. The Lomekwi 3 knappers, with a developing understanding of stone's fracture properties, combined core reduction with battering activities. Given the implications of the Lomekwi 3 assemblage for models aiming to converge environmental change, hominin evolution and technological origins, we propose for it the name 'Lomekwian', which predates the Oldowan by 700,000 years and marks a new beginning to the known archaeological record.

Conventional wisdom in human evolutionary studies has assumed that the origins of hominin sharp-edged stone tool production were linked to the emergence of the genus *Homo*^{1,2} in response to climate change and the spread of savannah grasslands^{3,4}. In 1964, fossils looking more like later *Homo* than australopithecines were discovered at Olduvai Gorge (Tanzania) in association with the earliest known stone tool culture, the Oldowan, and so were assigned to the new species: *Homo habilis* or 'handy man'. The premise was that our lineage alone took the cognitive leap of hitting stones together to strike off sharp flakes and that this was the foundation of our evolutionary success. Subsequent discoveries pushed back the date for the first Oldowan stone tools to 2.6 million years ago^{5,6} (Ma) and the earliest fossils attributable to early *Homo* to only 2.4–2.3 Ma^{7,8}, opening up the possibility of tool manufacture by hominins other than *Homo*⁹ before 2.6 Ma^{10–12}.

The earliest known artefacts from the sites of Gona (~2.6 Ma)^{6,12}, Hadar (2.36 ± 0.07 Ma¹³), and Omo (2.34 ± 0.04 Ma¹⁴) in Ethiopia, and especially Lokalalei 2C (2.34 ± 0.05 Ma¹⁵) in Kenya, demonstrate that these hominin knappers already had considerable abilities in terms of planning depth, manual dexterity and raw material selectivity^{14–19}. Cut-marked bones from Dikika, Ethiopia²⁰, dated at 3.39 Ma, has added to speculation on pre-2.6-Ma hominin stone tool use. It has been argued that percussive activities other than knapping, such as the pounding and/or battering of plant foods or bones, could have been critical components of an even earlier, as-yet-unrecognized, stage of hominin stone tool use^{21–25}. Any such artefacts may have gone unrecognized if they do not directly resemble known Oldowan lithics, occur at very low densities or were made of perishable materials¹⁰.

In 2011, the West Turkana Archaeological Project (WTAP) began an archaeological survey and excavation in the Lomekwi Member²⁶

(3.44–2.53 Ma) of the Nachukui Formation (west of Lake Turkana, northern Kenya; Fig. 1) to search for evidence of early hominin lithic behaviour. Several promising surface artefact concentrations and dispersed single finds were discovered. At the Lomekwi 3 archaeological site, 28 lithic artefacts were initially found lying on the surface or within a slope deposit, and one core was uncovered *in situ*. By the close of the subsequent 2012 field season, excavation at LOM3 had reached 13 m², revealing an additional 18 stone tools and 11 fossils *in situ* (Extended Data Table 1) within a horizon (approximately 80 cm) of indurated sandy-granular sediments stratified in a thick bed of fine silts (Fig. 2). A further 100 lithic artefacts and 22 fossil remains were collected from the surface immediately around the site along with two artefacts from the slope deposit (Extended Data Fig. 1). These finds occur in the same geographic and chronological range as the paratype of *Kenyanthropus platyops* (KNM-WT 38350)²⁷, other hominin fossils generally referred to cf. *K. platyops*²⁸, and one unpublished hominin tooth (KNM-WT 64060) found by WTAP in 2012 (Supplementary Information, part A and Supplementary Table 1).

Geochronological and palaeoenvironmental contexts

The chronological context of LOM3 derives from correlation with the Lomekwi Member of the Nachukui Formation²⁶ and radiometrically dated tuffs within it^{29,30}, as well as from magnetostratigraphy of the site and estimated sedimentation rates. The composite type section of the Lomekwi Member, 2–5 km east of LOM3, is bracketed by the α -Tulu Bor Tuff (3.44 ± 0.02 Ma) at the base and the Lokalalei Tuff (2.53 ± 0.02 Ma) at the top^{29,30}. Closer to LOM3, two new sections provide additional context. Section 1 (CSF 2011-1; ~46 m thick, located 1.44 to 1 km north of LOM3, Extended Data Fig. 2) includes

¹Turkana Basin Institute, Stony Brook University, Stony Brook, New York 11794-4364, USA. ²CNRS, UMR 7055, Préhistoire et Technologie, Université Paris Ouest Nanterre La Défense, 21 allée de l'Université, 92023 Nanterre Cedex, France. ³West Turkana Archaeological Project, P.O. Box 40658-00100, Ngara Rd, Nairobi, Kenya. ⁴Department of Anthropology and Center for Human Evolutionary Studies, Rutgers University, New Brunswick, New Jersey 08901, USA. ⁵Department of Earth and Planetary Sciences, Rutgers University, Piscataway, New Jersey 08854, USA. ⁶Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA. ⁷CNRS, UPR 2147, Dynamique de l'Évolution Humaine, 44 rue de l'Amiral Mouchez, 75014 Paris, France. ⁸CNRS, UMR 5199 PACEA, Université de Bordeaux, 33615 Pessac, France. ⁹Department of Sociology, Anthropology and Social Work, Seton Hall University, South Orange, New Jersey 07079, USA. ¹⁰Inrap, Centre Mixte de Recherche Archéologique, Domaine de Campagne, 24620 Campagne, France. ¹¹Inrap, 34-36 avenue Paul-Vaillant Couturier, 93120 La Courneuve, France. ¹²IPHEP, Institut de Paléoprimatologie, Paléontologie Humaine: Évolution et Paléoenvironnements, CNRS, UMR 7262, Université de Poitiers, Bât. B35 – TSA 51106, 6 rue Michel Brunet, 86073 Poitiers Cedex 9, France. ¹³Aix-Marseille Université, CNRS, MCC, UMR 7269, LAMPEA, 13094 Aix-en-Provence Cedex 2, France. ¹⁴National Museums of Kenya, Department of Earth Sciences, Archaeology Section, P.O. Box 40658-00100 Ngara Rd, Nairobi, Kenya.

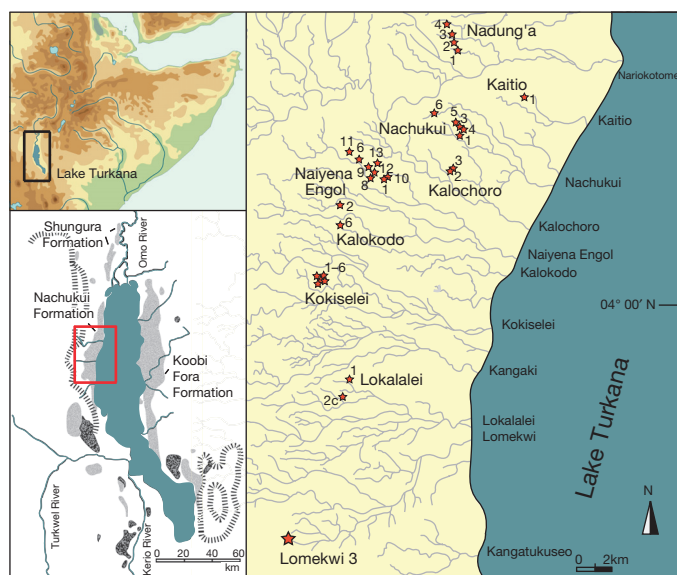


Figure 1 | Geographic location of the LOM3 site. Map showing relation of LOM3 to other West Turkana archaeological site complexes.

the α - and β -Tulu Bor Tuffs in the lower third (Supplementary Information, part B). Composite Section 2 (upper CSF-2012-9, ~44 m thick, located 0.4 km south of LOM3 and lower CSF-2011-2 located 0.28 km north of LOM3, Fig. 3a, b and Extended Data Fig. 2) includes at the base a lenticular tuff correlated geochemically with the Toroto Tuff in the Koobi Fora Formation where it outcrops 10–12 m above the α -Tulu Bor Tuff, and has been dated radiometrically to 3.31 ± 0.02 (refs 29, 30). Both the two Tulu Bor Tuffs in Section 1 and the Toroto Tuff in Section 2 occur in normal polarity magnetozones, corresponding to the early part of the Gauss Chron C2An (Fig. 3a and Supplementary Information, part C), while the overlying sediments at both sites are in reversed polarity zones as are the sediments encompassing the *in situ* artefacts at LOM3, 10 m above the Toroto Tuff (Fig. 3b). Thus, the artefacts were deposited after 3.31 ± 0.02 Ma during the Mammoth reverse subchron C2An.2r ($3.33\text{--}3.21$ Ma³¹). Based on extrapolation of sediment accumulation rates between the levels of the α -Tulu Bor and Toroto Tuffs and the onset of subchron C2An.2r, an age of 3.3 Ma is determined for LOM3 (Extended Data Fig. 3 and Supplementary Information, part C), which accords with previous interpretations of the antiquity of fossils from this locality^{27–30}.

Stable carbon isotopic analyses of pedogenic carbonate nodules located adjacent to and at LOM3 yielded a mean $\delta^{13}\text{C}_{\text{VPDB}}$ value of $-7.3 \pm 1.1\text{‰}$ (Extended Data Fig. 4 and Supplementary Information, part D), which indicates a mean fraction of woody cover (f_{wc}) of $47 \pm 9\%$ and positions the site within a woodland/bushland/thicket/shrubland environment³². Our results are comparable to paleosol $\delta^{13}\text{C}_{\text{VPDB}}$ values of other East African hominin environments between 3.2 and 3.4 Ma but significantly woodier than the 2.6 Ma artefact site at Gona, Ethiopia (Extended Data Figs 4b, c)^{32,33}. The associated fauna supports this interpretation (Supplementary Information, part E).

The Lomekwi 3 site

The LOM3 site is a low hill eroded into by a small ravine. The uppermost sediments encountered during excavation form a plaque of slope deposit which is a few centimetres thick (Fig. 2a). Under it, a series of interdigitated lenses of sands, granules and silts are found. They correspond to different facies of the same sedimentary environment related to the distal fan deposit in which the artefacts are preserved (Fig. 2c and Supplementary Information, part B). Sealed *in situ* in these Pliocene sediments (Extended Data Fig. 5), the LOM3 archaeological material is considered to be in a slightly re-distributed primary

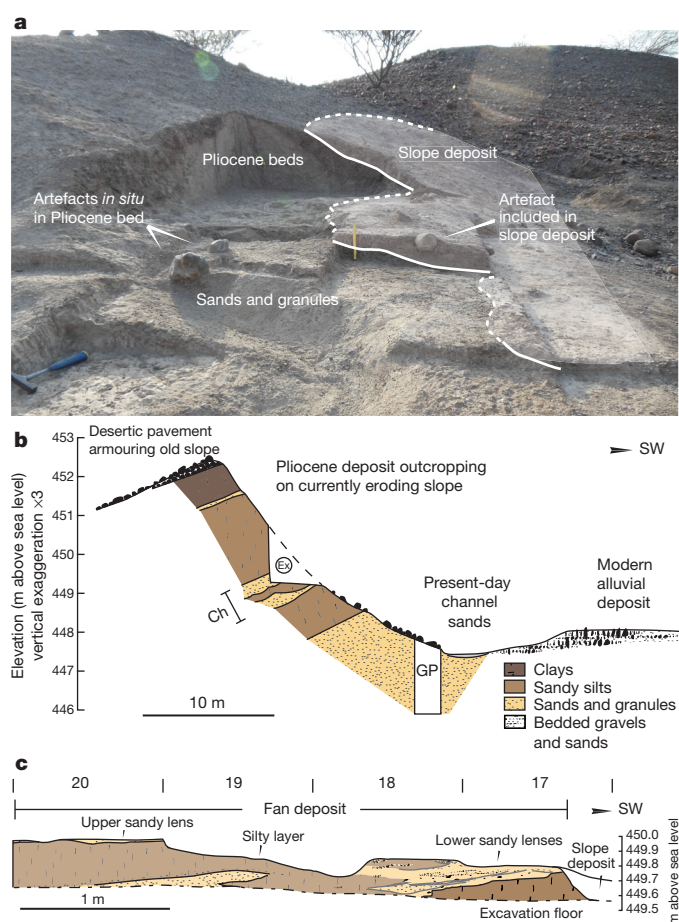


Figure 2 | LOM3 lithological context. a, View of the excavation, facing east, showing relationship between surface, slope deposit, and *in situ* contexts containing the artefacts and fossils. Scale in midground is 20 cm. Lower-leftmost artefact is the anvil LOM3-2012-K18-2, shown in Fig. 5a. b, Topographic profile and stratigraphic units at site level showing the excavation zone (Ex), the geological trench made at the base of the section (GP); the artefacts and fossils derive from a series of lenses of sand and granules making up a ~1 m thick bed (Ch). c, Section at the excavation along bands I and J (indicated by the black line in Extended Data Fig. 1a) showing the sediments which form the fan deposits containing the artefacts.

archaeological context based on the following observations: (1) artefacts of different sizes, ranging from ~1 cm wide flake fragments to very large worked cobbles and cores are present; (2) artefacts are larger and heavier than could be carried by the energy of the alluvial system that deposited the sediments (the maximal competence of the transport flow can be inferred by the coarsest fraction of the bed load deposited, that is, <4 cm diameter granules); (3) many excavated lithic pieces exhibit only slight abrasion, as reflected in the observation of arête and edge widths measuring $\leq 100 \mu\text{m}$. Moreover, although it is not possible at present to link all surface finds to the excavated context, the identification of a refit between a core recovered from the dense stratified deposit and one surface flake clearly shows that at least a portion of the surface material derives directly from the *in situ* level (Fig. 4a). More precise interpretation of site preservation is based on observations drawn from the excavation, with the most plausible possibilities limited to either good preservation of the site and most of the assemblage, or a slight redistribution in close proximity of the original activity location (Supplementary Information, part B).

Technology of the Lomekwi 3 stone tools

Based on the lithic material recovered in 2011 and 2012, the current total assemblage ($n = 149$ surface and *in situ* artefacts) incorporates

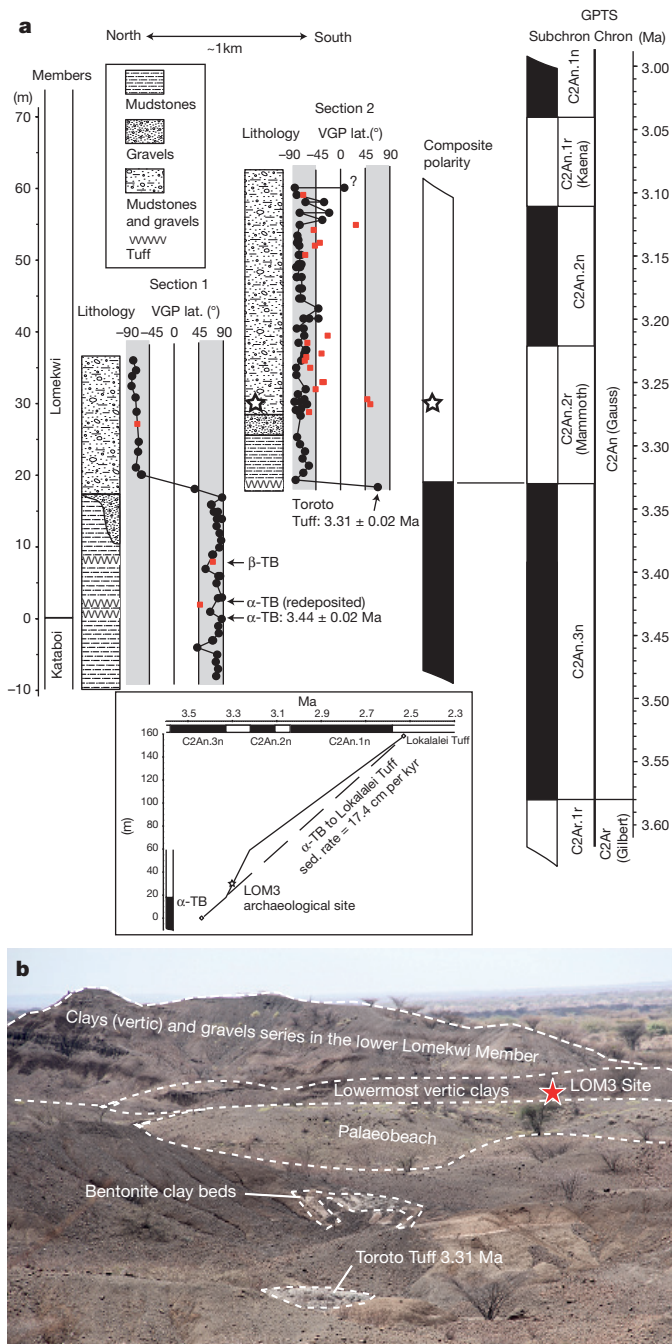


Figure 3 | Chronostratigraphic framework for LOM3.

a. Chronostratigraphic framework for LOM3 (star) with generalized stratigraphic columns and magnetostratigraphic alignment to the geomagnetic polarity time scale (GPTS) in context of dates of tuffaceous markers (± 1 s.d.) and stratigraphic nomenclature for Members of the Nachukui Formation^{26,30}. A linearly interpolated date of 3.3 Ma for the *in situ* stone tools is consistent with the site's magnetostratigraphic position within the reverse polarity interval that is correlated to reverse subchron C2An.2r (Mammoth Subchron) dated at 3.33–3.21 Ma³¹. **b.** Photograph facing north showing geographic and stratigraphic relationship between Toroto Tuff, paleobeach, and LOM3.

83 cores, 35 flakes (whole and broken), seven passive elements or potential anvils, seven percussors (whole, broken or potential), three worked cobbles, two split cobbles, and 12 artefacts grouped as indeterminate fragments or pieces lacking diagnostic attributes (Extended Data Table 1a).

Cores are made predominantly from heavy and large-sized cobbles or blocks of lava (mean of the cores: $167 \times 147.8 \times 108.8$ mm, 3.1 kg;

Extended Data Table 2). Basalts (34.90%) and phonolites (34.23%) are the dominant raw materials represented, followed by trachy-phonolite (23.49%; Extended Data Table 1b), all of which were available in local paleo-channels. Initial survey of a conglomerate source less than 100 m from the site shows that cobbles and blocks of all sizes were available locally, from which the largest were consistently selected. Most cores were flaked from one striking platform onto one single surface, resulting in several superposed and contiguous unidirectional removals (unifacial partial exploitation), sometimes along a longer part of the perimeter. A few specimens show unifacial partial exploitation by multidirectional removals, while others show bifacial flaking. Significant knapping accidents occurred during flaking, with numerous hinge and step flake terminations visible on cores (Fig. 4a), though more invasive and feather terminating flakes were also often successfully removed. In some cases, cores display a series of shorter (< 1 cm) contiguous small scars along a more limited portion of the platform edge, although it is not yet clear whether this results from the knapping techniques employed, or reflects the utilization of some artefacts in heavy-duty tasks.

To reconstruct more accurately the techniques and reduction strategies used to produce the LOM3 artefacts, an experimental program was undertaken to replicate the lithics found at the site from the same raw materials available locally at LOM3. Together with the technological analysis of the archaeological material, these replication experiments suggest that the LOM3 knappers were using techniques including passive hammer^{34,35} and/or bipolar³⁴ (Extended Data Fig. 6) that have to-date rarely been identified in the Oldowan^{21,22,36,37}. The average size and weight of the LOM3 cores (Extended Data Table 2) renders direct freehand percussion an arduous undertaking; however, it cannot be ruled out for some of the smaller cores.

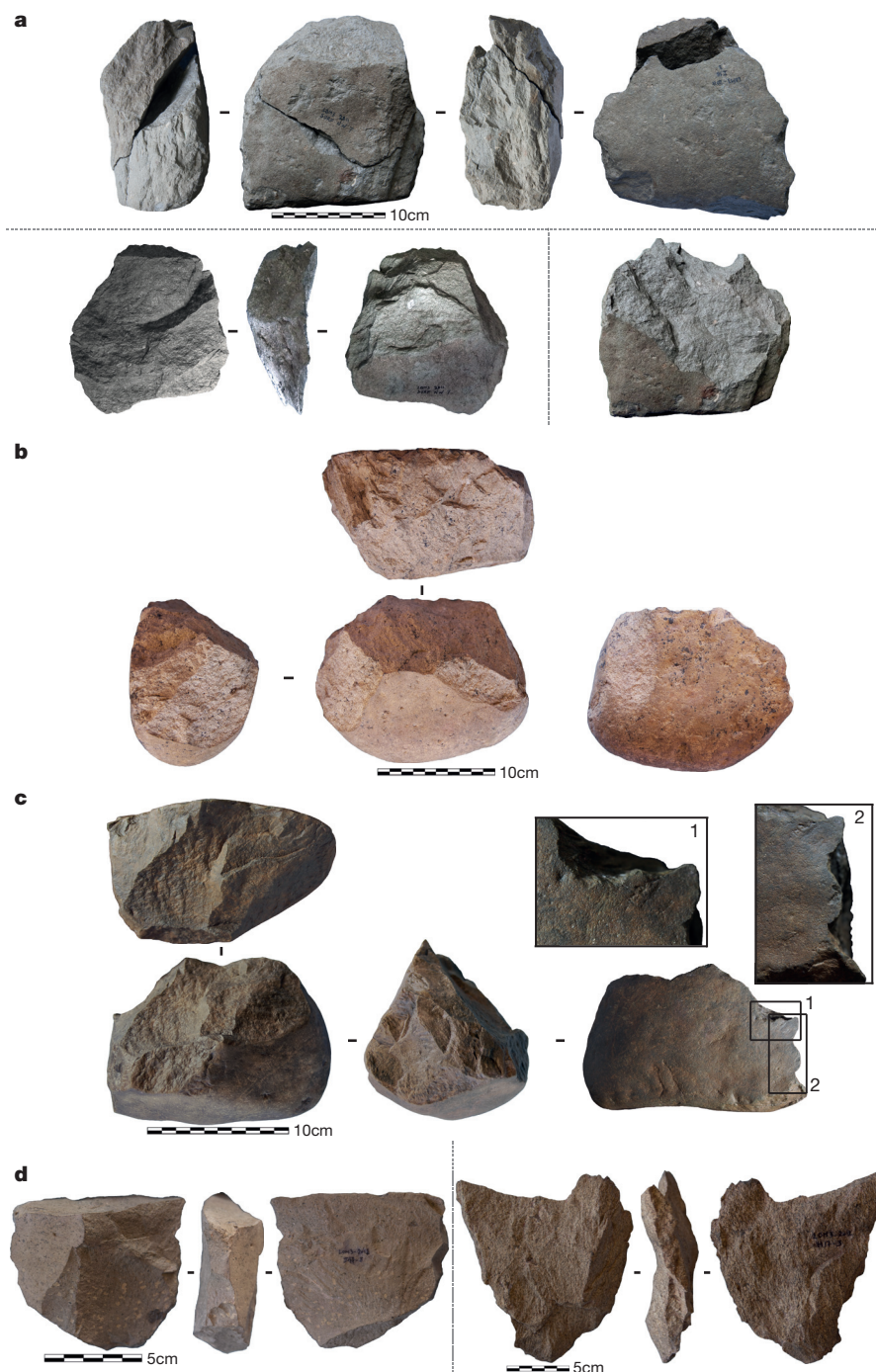
The technological features of flakes and flake fragments are clear, unequivocal and seen repeatedly, demonstrating that they were intentionally knapped from the cores. They range from 19 to 205 mm long (Fig. 5d and Extended Data Table 2) and frequently present cortex on their dorsal surfaces, sometimes on their striking platforms, or both. Three pieces in particular bear localized battered areas on their dorsal surfaces—including the specimen that refits onto the *in situ* core (Fig. 4a)—showing that blanks were sometimes used for percussive activities before flake removal and that at least some individual blocks were involved in several distinctively different modes of use.

The largest and heaviest (up to 15 kg) pieces in the assemblage were made on large blocks of basalt or coarse trachy-phonolite. They have flat natural surfaces that could enable their stabilization for use (Fig. 5a, b and Extended Data Fig. 7a). Comparisons with other described anvils from the Early Stone Age and experiments suggest these can be interpreted as anvils or passive elements^{38,39}. Three of these show a similar wear and fracture pattern. The largest piece exhibits along one lateral plane a series of divergent step fractures associated with crushing marks and an additional concentration of impact damage on one horizontal surface (Fig. 5a). The other two pieces have non-invasive step fractures along a greater or lesser portion of their high-angled intersecting surfaces (edges) that are associated with crushing and impact marks (Fig. 5b and Extended Data Fig. 7a). A further two cobbles show heavy battering marks concentrated on a convex area and are interpreted as passive elements. Seven medium-sized cobbles display battering marks and/or impact damage associated with fractured surfaces and are interpreted as hand-held percussors or active elements (Extended Data Figs 7b, c).

Discussion

LOM3 core and flake techno-morphology does not conform to any observed pattern resulting from accidental natural rock fracture. On the contrary, LOM3 cores and flakes bear all the techno-morphological characteristics of debitage products. Data reported on accidental flakes from chimpanzee nut-cracking sites⁴⁰ falls closer to the flake size spectrum observed at early Oldowan sites than to the size range of LOM3 flakes (Extended Data Table 2). LOM3 knappers were

Figure 4 | Photographs of selected LOM3 artefacts. **a**, *In situ* core (LOM3-2011-I16-3, 1.85 kg) and refitting surface flake (LOM3-2011 surf NW7, 650 g). Unifacial core, passive hammer and bipolar technique. Both the core and the flake display a series of dispersed percussion marks on cortex showing that percussive activities occurred before the removal of the flake, potentially indicating the block was used for different purposes. **b**, *In situ* unifacial core (LOM3-2012-H18-1, 3.45 kg), bipolar technique. See Extended Data Fig. 6b for more details. **c**, Unifacial core (LOM3-2012 surf 71, 1.84 kg), passive hammer technique. **d**, Flakes (LOM3-2012-J17-3 and LOM3-2012-H17-3) showing scars of previous removals on the dorsal face. See Supplementary Information part F for 3D scans of lithic artefacts.



able to deliver sufficient intentional force to repeatedly detach series of adjacent and superposed unidirectional flakes, sometimes invasive, and then to continue knapping either by laterally rotating the cores or by flipping them over for bifacial exploitation. However, though multiple flakes were successfully detached, the majority of flake scars terminate as hinge and step fractures. The precision of the percussive motion was also occasionally poorly controlled, as shown by repeated impact marks on core platforms caused by failed blows applied too far from the striking platform edge to induce fracture. LOM3 lithics (cores and flakes) are significantly larger in length, width, and thickness than those from OGS7, EG10 and EG12 at Gona, A.L. 894 at Hadar, and Omo 57 and Omo 123 in Ethiopia; Lokalalei 2C from West Turkana, Kenya; and DK and FLK Zinj from Olduvai Gorge in Tanzania (Extended Data Table 2). Furthermore, the LOM3 anvils and percussors are larger and heavier than those chosen for

nut-cracking by wild chimpanzees in Bossou⁴¹ (southeastern Guinea; Extended Data Table 3). The dimensions and the percussive-related features visible on the artefacts suggest the LOM3 hominins were combining core reduction and battering activities and may have used artefacts variously: as anvils, cores to produce flakes, and/or as pounding tools. The use of individual objects for several distinctive tasks reflects a degree of technological diversity both much older than previously acknowledged and different from the generally unipurpose stone tools used by primates^{24,25}. The arm and hand motions entailed in the two main modes of knapping suggested for the LOM3 assemblage, passive hammer and bipolar, are arguably more similar to those involved in the hammer-on-anvil technique chimpanzees and other primates use when engaged in nut cracking^{42–44} than to the direct freehand percussion evident in Oldowan assemblages. The likely prevalence of these two knapping techniques demonstrates

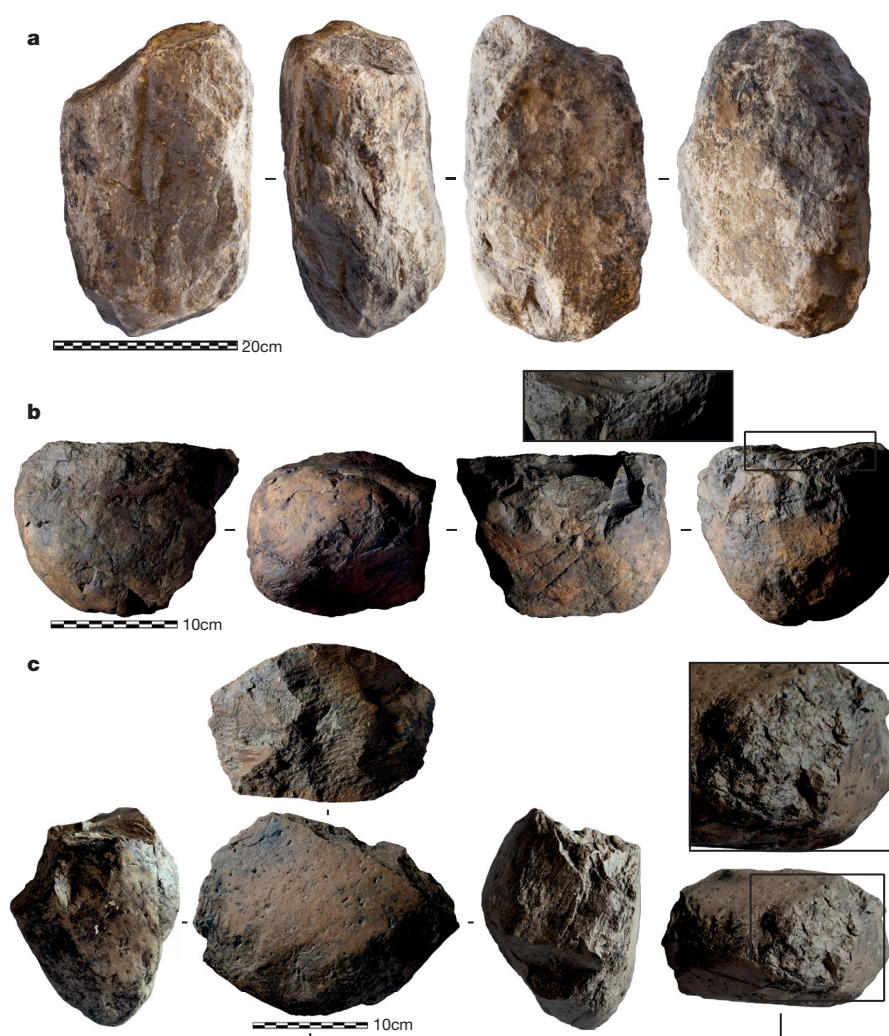


Figure 5 | Photographs of selected LOM3 artefacts. **a**, *In situ* passive element/anvil (LOM3-2012-K18-2, 12 kg). **b**, Passive element/anvil (LOM3-2012 surf 60, 4.9 kg). Both anvils **a** and **b** exhibit similar patterns of macroscopic wear consisting of superposed step fracturing in association with crushing and impacts marks. On **a**, damage is localized on a single lateral face, with battering marks present on one horizontal plane. On **b**, damage is distributed along a greater portion of the perimeter, but in this case no percussive marks are identifiable on the horizontal plane. In both cases, the intensity of the observed wear signature indicates a use in heavy-duty activities. **c**, Unifacial core (LOM3-2012 surf 90, 4.74 kg), bipolar technique and semi-peripheral exploitation. Inset shows crushing marks on the proximal surface of the cobble related to battering activities before or after the knapping of the core. See Supplementary Information part F for three-dimensional scans of lithic artefacts.

the central role that they might have played at the dawn of technology, as previously suggested^{21,22,36,37}.

LOM3 predates the oldest fossil specimens attributed to *Homo* in West Turkana at 2.34 ± 0.04 Ma⁷ by almost a million years; the only hominin species known to have been living in the West Turkana region at the time is *K. platyops*²⁷, while *Australopithecus afarensis* is found in the Lower Awash Valley at 3.39 Ma in association with cut-marked bones from Dikika²⁰. The LOM3 artefacts indicate that their makers' hand motor control must have been substantial and thus that reorganization and/or expansion of several regions of the cerebral cortex (for example, somatosensory, visual, premotor and motor cortex), cerebellum, and of the spinal tract could have occurred before 3.3 Ma. The functional morphology of the upper limb of Pliocene hominins (especially *A. afarensis*, the only species for which contemporaneous fossil hand and wrist elements are known), particularly in terms of adaptations for stone tool making, must be investigated further if this important milestone in human evolution is to be understood more fully (Supplementary Information, part A).

Critical questions relating to how the LOM3 assemblage compares with the previously known earliest hominin stone tool techno-complex, the Oldowan, remain. They are difficult to address because the term Oldowan has been defined differently since it was first employed in 1934 (refs 16, 45–47). The simplest defining characteristics of the Oldowan are that its knappers show the earliest evidence of a basic understanding of the conchoidal fracture mechanics of stone and were able to effectively strike flakes from cores, more often than not knapping using 'grammars of action'⁴⁸ and predominantly using the free-hand

knapping technique^{11,17}. The LOM3 knappers' understanding of stone fracture mechanics and grammars of action is clearly less developed than that reflected in early Oldowan assemblages and neither were they predominantly using free-hand technique. The LOM3 assemblage could represent a technological stage between a hypothetical pounding-oriented stone tool use by an earlier hominin and the flaking-oriented knapping behaviour of later, Oldowan toolmakers. The term 'Pre-Oldowan' has been suggested for modified stones if ever found in deposits older than 2.6 Ma, especially if they are different in terms of knapping skill from the Oldowan *sensu stricto*⁴⁹ (this is not to be confused with previous uses of the same term by some authors to describe the early Oldowan period between 2.6–2 Ma⁵⁰). The LOM3 assemblage may therefore concord with such a premise. We assert, however, that the technological and morphological differences between the LOM3 and early Oldowan assemblages are significant enough that amalgamating them would mask important behavioural and cognitive changes occurring among hominins over a nearly 2-million-year timespan. A separate name for the LOM3 assemblage is therefore warranted. Given the paradigmatic shift that LOM3 portends for models that aim to converge environmental change, hominin evolution and technological origins, the name Lomekwian is proposed. In any scenario, the LOM3 stone tools mark a new beginning to the known archaeological record, now shown to be more than 700,000 years older than previously thought.

Note added in proof: The recently described LD 350-1 partial mandible from Ethiopia now provides the earliest evidence of the genus *Homo* at 2.8 Ma (Villmoare, B. *et al.* Early *Homo* at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia. *Science*, 347, 1352–1355). The LOM3 artefacts

still predate the known origins of *Homo* by half a million years and the question of what hominin species made them remains.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 November 2012; accepted 13 April 2015.

- Leakey, L. S. B., Tobias, P. V. & Napier, J. R. A new species of the genus *Homo* from Olduvai Gorge. *Nature* **202**, 7–9 (1964).
- Harris, J. W. K. Cultural beginnings: Plio-Pleistocene archaeological occurrences from the Afar Rift, Ethiopia. *Afr. Archaeol. Rev.* **1**, 3–31 (1983).
- Quinn, R. L. *et al.* Pedogenic carbonate stable isotopic evidence for wooded habitat preference of early Pleistocene tool makers in the Turkana Basin. *J. Hum. Evol.* **65**, 65–78 (2013).
- Bobé, R. & Behrensmeyer, A. K. The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus *Homo*. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **207**, 399–420 (2004).
- Roche, H. & Tiercelin, J.-J. Découverte d'une industrie lithique ancienne *in situ* dans la formation d'Hadar, Afar central, Ethiopie. *C. R. Acad. Sci. Paris D* **284**, 1871–1874 (1977).
- Semaw, S. *et al.* 2.5-million-year-old stone tools from Gona, Ethiopia. *Nature* **385**, 333–336 (1997).
- Prat, S. *et al.* First occurrence of early *Homo* in the Nachukui Formation (West Turkana, Kenya) at 2.3–2.4 Myr. *J. Hum. Evol.* **49**, 230–240 (2005).
- Kimbel, W. H. *et al.* Late Pliocene *Homo* and Oldowan tools from the Hadar formation (Kada Hadar member), Ethiopia. *J. Hum. Evol.* **31**, 549–561 (1996).
- Antón, S. C., Potts, R. & Aiello, L. C. Evolution of early *Homo*: An integrated biological perspective. *Science* **345**, 1236828 (2014).
- Panger, M., Brooks, A. S., Richmond, B. G. & Wood, B. Older than the Oldowan? Rethinking the emergence of hominin tool use. *Evol. Anthropol.* **11**, 235–245 (2002).
- Roche, H., Blumenschine, R. J. & Shea, J. J. In *The First Humans — Origin and Early Evolution of the Genus Homo* (eds Grine, F. E., Fleagle, J. G. & Leakey, R. E.) 135–147 (Springer, 2009).
- Semaw, S. *et al.* 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *J. Hum. Evol.* **45**, 169–177 (2003).
- Campisano, C. J. Geological summary of the Busidima Formation (Plio-Pleistocene) at the Hadar paleoanthropological site, Afar Depression, Ethiopia. *J. Hum. Evol.* **62**, 338–352 (2012).
- de la Torre, I. Omo revisited: evaluating the technological skills of Pliocene hominids. *Curr. Anthropol.* **45**, 439–465 (2004).
- Roche, H. *et al.* Early hominid stone tool production and technical skill 2.34 Myr ago in West Turkana, Kenya. *Nature* **399**, 57–60 (1999).
- Delagnes, A. & Roche, H. Late Pliocene hominid knapping skills: the case of Lokalalei 2C, West Turkana, Kenya. *J. Hum. Evol.* **48**, 435–472 (2005).
- Stout, D., Semaw, S., Rogers, M. J. & Cauche, D. Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *J. Hum. Evol.* **58**, 474–491 (2010).
- Harmand, S. in *Interdisciplinary Approaches to the Oldowan* (eds Hovers, E. & Braun, D. R.) 85–97 (Springer, 2009).
- Goldman-Neuman, T. & Hovers, E. Raw material selectivity in Late Pliocene Oldowan sites in the Makaamitalu Basin, Hadar, Ethiopia. *J. Hum. Evol.* **62**, 353–366 (2012).
- McPherron, S. P. *et al.* Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* **466**, 857–860 (2010).
- Mora, R. & de la Torre, I. Percussion tools in Olduvai Beds I and II (Tanzania): implications for early human activities. *J. Anthropol. Archaeol.* **24**, 179–192 (2005).
- Diez-Martín, F., Sánchez Yustos, P., Domínguez-Rodrigo, M., Mabulla, A. Z. P. & Barba, R. Were Olduvai hominins making butchering or battering tools? Analysis of a recently excavated lithic assemblage from BK (Bed II, Olduvai Gorge, Tanzania). *J. Anthropol. Archaeol.* **28**, 274–289 (2009).
- Blumenschine, R. J. & Selvaggio, M. M. Percussion marks on bone surfaces as a new diagnostic of hominid behaviour. *Nature* **333**, 763–765 (1988).
- Marchant, L. & McGrew, W. in *Stone Knapping: the Necessary Conditions for a Uniquely Hominin Behavior* (eds Roux, V. & Brill, B.) 341–350 (Cambridge McDonald Institute, 2005).
- Carvalho, S., Cunha, E., Sousa, C. & Matsuzawa, T. *Chânes opératoires* and resource-exploitation strategies in chimpanzee (*Pan troglodytes*) nut cracking. *J. Hum. Evol.* **55**, 148–163 (2008).
- Harris, J. M., Brown, F. H. & Leakey, M. G. Stratigraphy and paleontology of Pliocene and Pleistocene localities west of Lake Turkana, Kenya. *Contr. Sci. Nat. Mus. Los Angeles* **399**, 1–128 (1988).
- Leakey, M. G. *et al.* New hominin genus from eastern Africa shows diverse middle Pliocene lineages. *Nature* **410**, 433–440 (2001).
- Wood, B. & Leakey, M. G. The Omo-Turkana Basin fossil hominins and their contribution to our understanding of human evolution in Africa. *Evol. Anthropol.* **20**, 264–292 (2011).
- McDougall, I. & Brown, F. H. Geochronology of the pre-KBS Tuff sequence, Turkana Basin. *J. Geol. Soc. Lond.* **165**, 549–562 (2008).
- McDougall, I. *et al.* New single crystal $^{40}\text{Ar}/^{39}\text{Ar}$ ages improve time scale for deposition of the Omo Group, Omo-Turkana Basin, East Africa. *J. Geol. Soc. Lond.* **169**, 213–226 (2012).
- Lourens, L. J., Hilgen, F. J., Laskar, J., Shackleton, N. J. & Wilson, D. *The Neogene Period. A Geological Time Scale 2004* (eds Gradstein, F. M., Ogg, J. G. & Smith, A. G.) 409–440 (Cambridge University Press, 2004).
- Cerling, T. E. *et al.* Woody cover and hominin environments in the past 6 million years. *Nature* **476**, 51–56 (2011).
- Quade, J. *et al.* Paleoenvironments of the earliest stone toolmakers, Gona, Ethiopia. *Geol. Soc. Am. Bull.* **116**, 1529–1544 (2004).
- Crabtree, D. E. An introduction to flintworking. Part 1. An introduction to the technology of stone tools. *Occasional Papers of the Idaho State University Museum*, **28**, (Idaho State Univ. Museum, 1972).
- Mourre, V., Jarry, M., Colonge, D. & Lelouvier, L.-A. Le débitage sur enclume aux Bosses (Lamagdalaine, Lot, France). *Paleo* (special issue) 49–62 (2010).
- Diez-Martín, F. *et al.* New insights into hominin lithic activities at FLK North Bed I, Olduvai Gorge, Tanzania. *Quat. Res.* **74**, 376–387 (2010).
- de la Torre, I. & Mora, R. A technological analysis of non-flaked stone tools in Olduvai Beds I & II. Stressing the relevance of percussion activities in the African Lower Pleistocene. *Paleo* (special issue) 13–34 (2010).
- Alimen, M. H. Enclumes (percuteurs dormants) associées à l'Acheuléen supérieur de l'Ougartien (Oued Farès, Sahara occidental). *Bull. Soc. Préhist. Fr.* **60**, 43–47 (1963).
- Leakey, M. D. *Olduvai Gorge, Vol. 3. Excavations in Beds I and II 1960–1963* (Cambridge Univ. Press, 1971).
- Mercader, J. *et al.* 4,300-year-old chimpanzee sites and the origins of percussive stone technology. *Proc. Natl Acad. Sci. USA* **104**, 3043–3048 (2007).
- Sakura, O. & Matsuzawa, T. Flexibility of wild chimpanzee nut-cracking behavior using stone hammers and anvils: an experimental analysis. *Ethology* **87**, 237–248 (1991).
- Haslam, M. *et al.* Primate archaeology. *Nature* **460**, 339–344 (2009).
- Visalberghi, E., Haslam, M., Spagnoletti, N. & Frigaszy, D. Use of stone hammer tools and anvils by bearded capuchin monkeys over time and space: construction of an archeological record of tool use. *J. Archaeol. Sci.* **40**, 3222–3232 (2013).
- Matsuzawa, T. in *Great Ape Societies* (eds McGrew, W. *et al.*) 196–209 (Cambridge Univ. Press, 1996).
- Leakey, L. S. B. in *Essays Presented to C. G. Seligman* (eds Evans-Pritchard, E. E., Firth, R. Malinowski, B. & Schapera, I.) 143–146 (K. Paul, Trench, Trubner & Co, 1934).
- de la Torre, I. The origins of stone tool technology in Africa: a historical perspective. *Phil. Trans. R. Soc. B.* **366**, 1028–1037 (2011).
- Shea, J. Lithic modes A-I: a new framework for describing global-scale variation in stone tool technology illustrated with evidence from the East Mediterranean Levant. *J. Archaeol. Method Theory* **20**, 151–186 (2013).
- Hovers, E. in *Origins of Human Innovation and Creativity* (ed. Elias, S.) 51–68 (Elsevier, 2012).
- Toth, N., Schick, K. & Semaw, S. in *The Oldowan: Case Studies into the Earliest Stone Age* (eds Toth, N. & Schick, K. D.) 155–222 (Stone Age Institute Press, 2006).
- Piperno, M. in *Hominidae: Proc. 2nd Intl Congr. Human Paleontol. 1987* 189–195 (Jaca Books, 1989).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the office of the President of Kenya, the Ministry of Education, Science and Technology, the National Council for Science and Technology (NCST/RCD/12B/012/25) and the National Museums of Kenya for permission to conduct research. Funding was provided by the French Ministry of Foreign Affairs (N°681/DGM/ATT/RECH, N°986/DGM/DPR/PRG), the French National Research Agency (ANR-12-CULT-0006), the *Fondation Fyssen*, the National Geographic Society (Expeditions Council #EC0569-12), the Rutgers University Research Council and Center for Human Evolutionary Studies, and INTM Indigo Group France. We thank the Turkana Basin Institute and Total Kenya Limited for logistical support and the GeoEye Foundation for satellite imagery; the Turkana communities from Nariokotome, Kokiselei and Katiko for field assistance, and the 2011–12 WTAP team, S. Kahinju, P. Egoan, L. P. Martin, D. Massika, B. K. Muluwa S. M. Musyoka, A. Mutisiya, J. Mwambua, F. M. Wambua, M. Terrade, A. Weiss, R. Benitez, S. Feibel, M. Leakey and F. Spoor supplied information on hominin fossils, and I. de la Torre and E. Hovers provided lithic assemblage data. We are very grateful to A. Brooks, I. de la Torre, J. Shea, R. Klein and M. Leakey for comments on earlier drafts. We also thank the Zoller & Fröhlich GmbH company, Ch. Fröhlich and M. Reinköster, Autodesk and Faro (T. O'Mahoney, K. Almeida Warren and T. Gichunge) for technical support with scanning and J. P. Chirey for photographic assistance.

Author Contributions S.H. and J.E.L. directed field research and co-wrote the overall paper. C.S.F., C.J.L., A.L. and X.B. recorded sedimentological and stratigraphic data, conducted geological mapping, and wrote sections of the paper. C.S.F. interpreted tephra data. C.J.L. interpreted paleomagnetic data. S.P., J.-Ph.B., S.L., C.K. and L.L. conducted paleontological survey. S.P., J.-Ph.B. and L.L. analysed and interpreted fossil material. L.L. directed scanning of artefacts. S.P. laser scanned artefacts and excavation surfaces, and wrote sections of the paper. R.L.Q. interpreted isotopic data and wrote sections of the paper. C.S.F., C.J.L., R.L.Q., R.A.M., J.D.W. and D.V.K. analysed geological samples. G.D. developed protocols for tool replication experiments and wrote sections of the paper. S.H., H.R., N.T., M.B., S.C., S.L. and C.K. conducted archaeological survey and excavation. S.H., H.R., A.A., N.T. and M.B. analysed and interpreted lithic material and wrote sections of the paper. M.B. performed lithic replication experiments. S.C. provided spatial data. S.L. discovered the LOM3 site.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.H. (sonia.harmand@stonybrook.edu) or J.L. (jason.lewis@stonybrook.edu).

METHODS

Paleomagnetic analyses. All samples from the Lomekwi outcrops were collected from fresh surfaces uncovered by digging into the exposures for at least 20 cm. Before each hand-cut block was extracted, *in situ* azimuths and dips were recorded on a sample using a compass-inclinometer. Samples were taken typically at nominal 1 m vertical stratigraphic intervals, or as the distribution of fine-grained strata allowed. Two sections were sampled, separated from each other by about 1 km north to south across the landscape (Extended Data Fig. 2). Overlapping Sections 1 and 2 (Fig. 3a) are each composed of a coarsening upward succession of mudstones abruptly overlain by gravels and followed by a thick unit of gravels and mudstones, which likely records a lacustrine regression and the emplacement of a prograding alluvial fan. Inset in Fig. 3a shows stratigraphic thickness of composite section plotted against key chronostratigraphic levels (α -Tulu Bor (α -TB), 3.44 ± 0.02 ; Toroto Tuff, 3.31 ± 0.02 Ma; C2An.3n/2r boundary, 3.33 Ma³¹; Lokalei Tuff, 2.53 ± 0.02).

At Section 1 (Fig. 3a), sampling began at about 10 m below the lowermost stratigraphic level of the α -Tulu Bor Tuff. Sampling continued upwardly from the α -Tulu Bor Tuff for another 35 m, for a total of ~45 m sampled. At Section 2 (Fig. 3a), sampling commenced at the Toroto Tuff. Sampling started upwardly from the Toroto Tuff for about 10 m to the level of the archaeological horizon at LOM3, and then proceeded upwardly for another 35 m for a total sampled stratigraphic thickness of about 45 m at Section 2.

For laboratory analyses, samples were cut into standard cube-shape specimens (~10 cc) using a lapidary saw and sandpaper. All magnetic remanence measurements were made with a 2G DCSQUID rock magnetometer in the shielded room at the Paleomagnetism Laboratory of Lamont-Doherty Earth Observatory (Columbia University). The natural remanent magnetization (NRM) of a specimen was subjected to progressive Thermal Demagnetization (TD) using 14 to 17 steps at 100, 50 and 25 °C increments in the temperature range of 100–700 °C. Data from consecutive high-temperature steps were used for principal component analysis (PCA³¹) to fit least-square lines tied to the origin for the final demagnetization trajectories defining the characteristic remanent magnetization (ChRM) as revealed on orthogonal projection plots (Extended Data Fig. 3a). Magnetic susceptibility values were determined with a Bartington MS2B instrument for each specimen initially and after each TD heating step to monitor any laboratory-induced magnetochemical alteration. The virtual geomagnetic pole (VGP) latitude corresponding to the ChRM direction was used to determine the magnetostratigraphic polarity sequence. In Fig. 3a, filled black circles joined by lines (isolated red squares) denote accepted (rejected) data with maximum angular deviation (MAD) values <15° (>15°) from principal component analyses. Characteristic remanent magnetizations were isolated after the removal of a pervasive normal polarity overprint unblocked by a TD range of 600–670 °C for the coarse alluvial fan strata (essentially all of Section 2 above the Toroto Tuff) and a TD range of 400–550 °C for the finer strata (for example, mudstones from the lower part of Section 1).

Pedogenic carbonate stable carbon isotopic analysis. Sedimentological field analysis identified eleven paleosols with discernible preserved B_K horizons. Ten paleosols were sampled from Section 2011-1, and one from 2011-2 (Extended Data Fig. 2). Carbonate nodules were extracted from paleosols >30 cm below the contact with overlying stratum with vertic features within peds showing slickensided surfaces. Twenty-four cross-sectioned nodules (five from one paleosol at LOM3, 2011-2) were sampled with a 0.5 mm carbide drill bit (Foredom Series) and loaded into v-vials for single acid baths (multi-prep device). Forty-seven isotopic analyses were conducted on a Micromass Optima mass spectrometer in the Department of Earth and Planetary Sciences at Rutgers University. Samples were reacted at 90 °C in 100% phosphoric acid for 13 min. $\delta^{13}\text{C}_{\text{VPDB}}$ values are reported in the standard per mil (‰) notation: $= (R_{\text{sample}}/R_{\text{standard}} - 1) \times 1000$, relative to Vienna-Pee Dee Belemnite through analysis of laboratory standard NBS-19 (Extended Data Fig. 4). Analytical error is $\pm 0.05\text{‰}$. Using methods of ref. 32, we subtracted 14‰ from the $\delta^{13}\text{C}_{\text{VPDB}}$ values of pedogenic carbonate to convert to the isotopic equivalent of organic carbon ($\delta^{13}\text{C}_{\text{om}}$) and used the equation: $f_{\text{wc}} = \{\sin[-1.06688 - 0.08538(\delta^{13}\text{C}_{\text{om}})]\}^2$ to generate estimates of fraction woody canopy cover for classification into UNESCO categories of African vegetation. Categories were taken from White⁵² and have the following $\delta^{13}\text{C}_{\text{VPDB}}$ value ranges of pedogenic carbonates³²: (1) forest: continuous stand of trees at least 10-m tall with interlocking crowns with greater than 80% woody cover ($\delta^{13}\text{C}_{\text{VPDB}}$: > -11.5

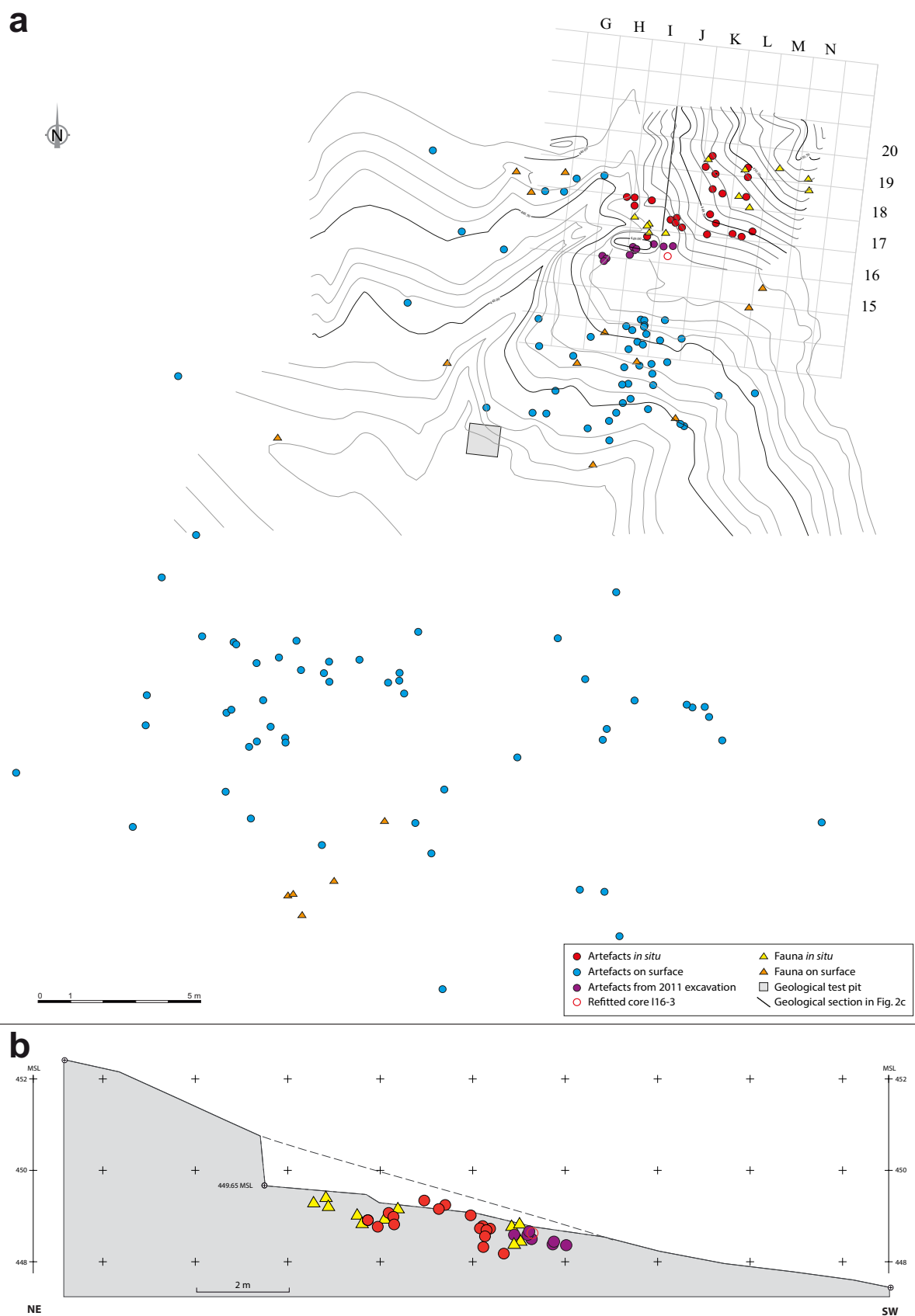
‰), (2) woodland/bushland/thicket/shrubland: woodland is an open stand of trees at least 8 m tall with woody cover >40% and a field layer dominated by grasses; bushland is an open stand of bushes between 3 m and 8 m tall with woody cover >40%; thicket is a closed stand of bushes and climbers between 3 m and 8 m tall; shrubland is an open or closed stand of shrubs up to 2 m tall ($\delta^{13}\text{C}_{\text{VPDB}}$: -11.5 to -6.5‰), (3) wooded grassland: land covered with grassland and has 10–40% tree or shrub cover ($\delta^{13}\text{C}_{\text{VPDB}}$: -6.5 to -2.3 ‰) and (4) grassland: land covered with herbaceous plants with less than 10% tree and shrub cover ($\delta^{13}\text{C}_{\text{VPDB}}$: < -2.3 ‰). We also calculated percent C₄ biomass using a simple linear mixing model assuming -12‰ and -26‰ as the C₄ and C₃ end members, respectively⁵³.

Site scanning. To document the uncovering of the *in situ* artefacts and fossils during the excavation, we took frequent 3D scans of the surface of individual squares with the OptiNum RE handheld device (manufactured by Noomeo Products, France) with a maximum spatial resolution of 300 µm. Additionally, thanks to a collaboration between Zoller & Fröhlich GmbH and Autodesk, we had access to a recently developed high-resolution industrial 3D scanner operated by M. Reinköster. This scanner was able to scan the entire site, registering 500,000 3D points each second, with a spatial resolution of <3,000 µm, and recording for several minutes continuously. After the laser scan a 3D photo was taken which can be draped around the scan. The site was scanned in this manner after each day of excavation. In this way, a high-resolution 3D digital model can be created for the entire site, and individual squares, showing the evolution of the excavation and the original context and gradual uncovering of the *in situ* artefacts and fossils.

Stone tool scanning. A representative sample of the LOM3 artefacts were scanned at the National Museums of Kenya and the Turkana Basin Institute facility in Turkwel, using a LMI Technologies R3 Advance portable structured light scanner (LMI Technologies, Vancouver, Canada), calibrated to the size of the objects in question, with the calibration grids being accurate to 50 µm. For colour texture overlay, a Canon 600D/Rebel T3i SLR digital camera was also calibrated with the scanner and images from this formed the base of the colour texture. The textured files are saved in .obj format and non textured files (for 3D printing or similar purposes) are saved in .stl format. These scans and 3-D digital models are available at (<http://africanfossils.org/search>).

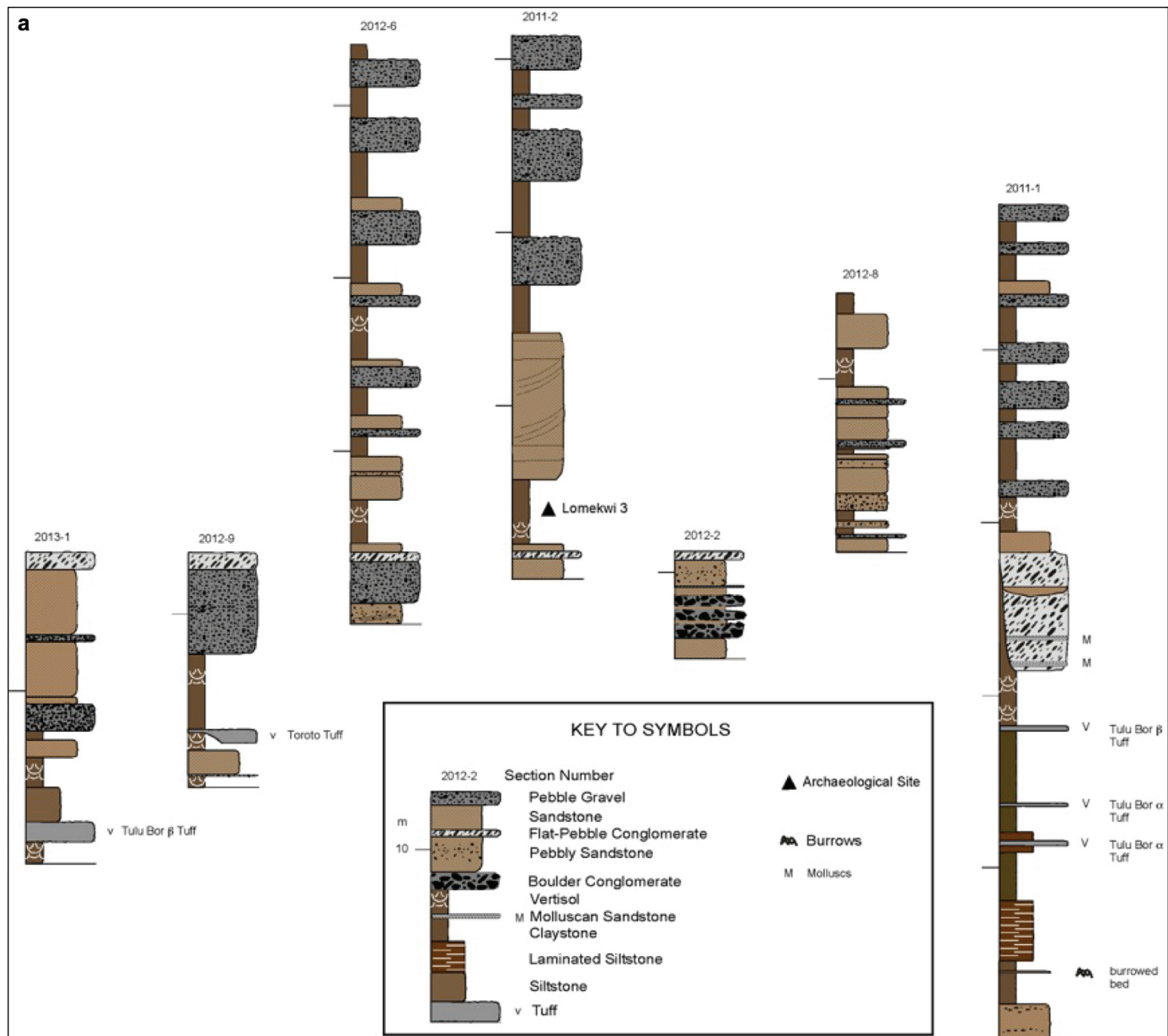
Sample size. No statistical methods were used to predetermine sample size

51. Kirschvink, J. L. The least-squares line and plane and the analysis of palaeomagnetic data. *Geophys. J. Int.* **62**, 699–718 (1980).
52. White, F. The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. *UNESCO. Nat. Resour. Res.* **20**, 1–356 (1983).
53. Fox, D. L. & Koch, P. L. Carbon and oxygen isotopic variability in Neogene paleosol carbonates: constraints on the evolution of the C₄-grasslands of the Great Plains, USA. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **207**, 305–329 (2004).
54. Levin, N. E., Quade, J., Simpson, S. W., Semaw, S. & Rogers, M. J. Isotopic evidence for Plio-Pleistocene environmental change at Gona, Ethiopia. *Earth Planet. Sci. Lett.* **219**, 93–110 (2004).
55. Levin, N. E. Compilation of East Africa soil carbonate stable isotope data. *Integrated Earth Data Applications* <http://dx.doi.org/10.1594/IEDA/100231> (2013).
56. Cerling, T. E., Bowman, J. R. & O'Neil, J. R. An isotopic study of a fluvial-lacustrine sequence: the Plio-Pleistocene Koobi Fora sequence, East Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **63**, 335–356 (1988).
57. Levin, N. E., Brown, F. H., Behrensmeyer, A. K., Bobe, R. & Cerling, T. E. Paleosol carbonates from the Omo Group: isotopic records of local and regional environmental change in East Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **307**, 75–89 (2011). CrossRef.
58. Wynn, J. G. Influence of Plio-Pleistocene aridification on human evolution: evidence from paleosols from the Turkana Basin, Kenya. *Am. J. Phys. Anthropol.* **123**, 106–118 (2004).
59. Kingston, J. D. *Stable isotopic evidence for hominid paleoenvironments in East Africa*. Ph.D. Thesis, Harvard Univ. (1992).
60. Aronson, J. L., Hailemichael, M. & Savin, S. M. Hominid environments at Hadar from paleosol studies in a framework of Ethiopian climate change. *J. Hum. Evol.* **55**, 532–550 (2008).
61. Wynn, J. G. et al. Geological and palaeontological context of a Pliocene juvenile hominin at Dikika, Ethiopia. *Nature* **443**, 332–336 (2006).
62. Semaw, S., Rogers, M. J. & Stout, D. In *The Cutting Edge: New Approaches to the Archaeology of Human Origins* (eds Schick, K. D. & Toth, N.) 211–246 (Stone Age Institute Press, 2009).
63. Hovers, E. In *The Cutting Edge: New Approaches to the Archaeology of Human Origins* (eds Schick, K. D. & Toth, N.) 137–150 (Stone Age Institute Press, 2009).
64. de la Torre, I. & Mora, R. *Technological Strategies in the Lower Pleistocene at Olduvai Beds I & II*. ERAUL 112 (2005).



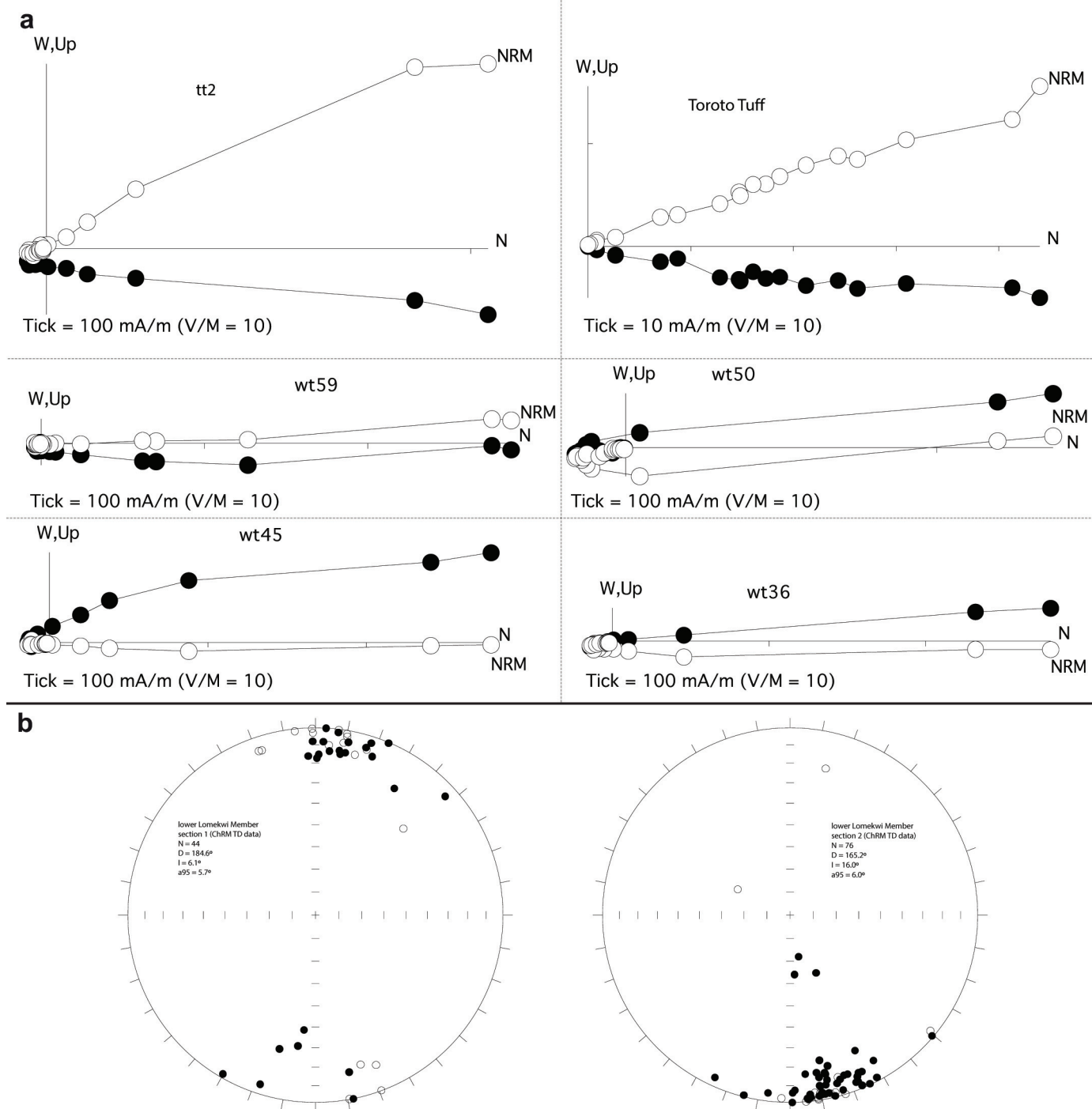
Extended Data Figure 1 | Map and schematic section at LOM3. **a**, Map showing *xy* coordinates of artefacts and fossils recovered *in situ* and from the surface at the site in 2011 and 2012. **b**, Schematic section showing vertical

distribution of *in situ* artefacts and those located in the slope deposit at the excavation. Key is the same for both figures.

**b**

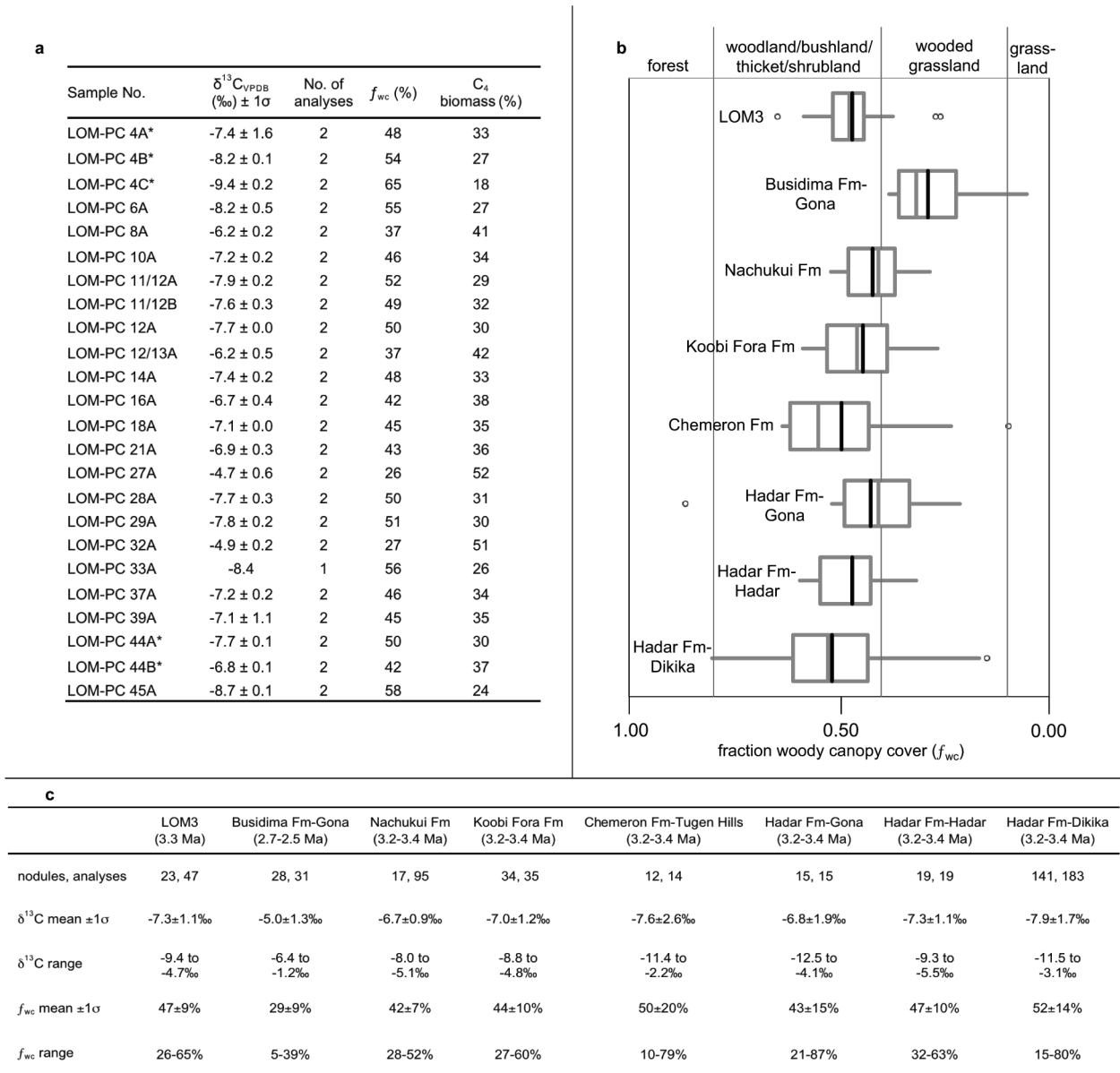
	N Lat (d mm.mmm)	E Long (dd mm.mmm)
base 2011-1	3 52.441	35 45.201
top 2011-1	3 52.219	35 45.325
base 2011-2	3 51.806	35 45.183
top 2011-2	3 51.814	35 45.292
base 2012-2	3 51.627	35 45.228
top 2012-2	3 51.662	35 45.251
base 2012-6	3 51.875	35 44.932
top 2012-6	3 51.969	35 44.971
base 2012-8	3 51.679	35 45.268
top 2012-8	3 51.699	35 45.282
base 2012-9	3 51.473	35 45.205
top 2012-9	3 51.512	35 45.181
base 2013-1	3 51.485	35 44.973
top 2013-1	3 51.511	35 45.003

Extended Data Figure 2 | Geology of the LOM3 site. a, Stratigraphic sections around LOM3 (locations in b), showing relationship of site to marker tuffs and lithofacies. Sections aligned relative to top of flat-pebble conglomerate unit. b, GPS coordinates of stratigraphic sections (WGS84 datum).



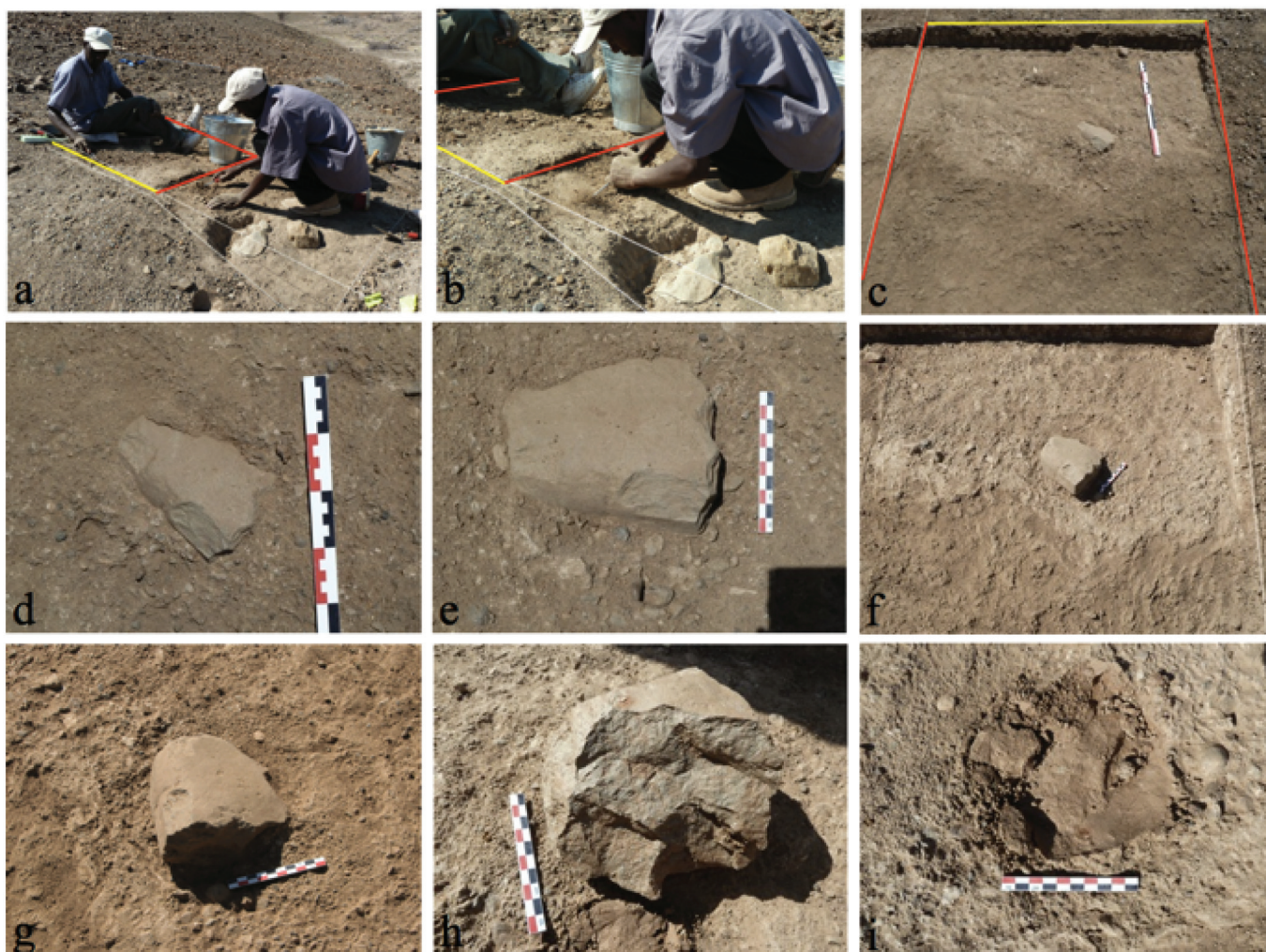
Extended Data Figure 3 | Paleomagnetic data. **a**, Representative vector end-point plots of natural remanent magnetism thermal demagnetization data from specimen Toroto Tuff, tt2, wt59, wt50, wt45, wt36. Open and closed symbols represent the vertical and horizontal projections, respectively, in bedding coordinates. TD treatment steps: NRM, 100°, 150°, 200°, 250°, 300°, 350°, 400°, 450°, 475°, 500°, 525°, 550°, 575°, 600°, 625°, 650°, 660°, 670°, 675°, 680°, 690°, and 700°. V/M = 10 denotes a ~10 cc cubic specimen. **b**, Equal-area

projections for Section 1 (left) and Section 2 (right) of the lower Lomekwi Member (see Fig. 3a). Open and closed symbols are projected onto the upper and lower hemisphere, respectively, in bedding coordinates. Plotted are ChRM sample-mean directions for accepted samples only (that is, those with MAD values <15°). Overall mean directions were calculated after inverting the northerly (normal) directions to common southerly (reverse) polarity.



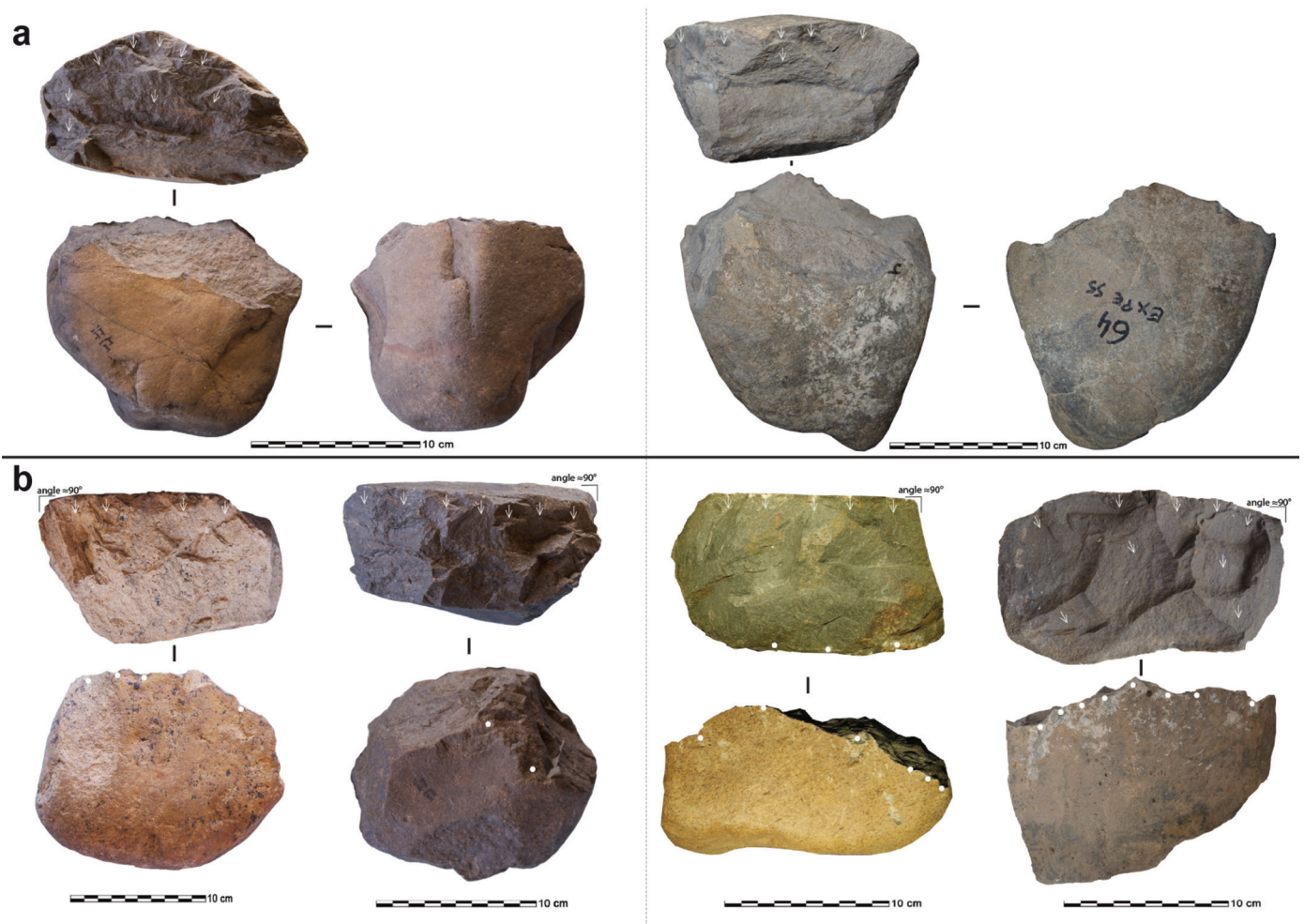
Extended Data Figure 4 | Paleoenvironmental reconstruction through pedogenic carbonate stable carbon isotopic analysis. **a**, LOM3 paleosol $\delta^{13}\text{C}_{\text{VPDB}}$ values (‰) $\pm 1\sigma$, number of analyses, fraction woody canopy cover (f_{wc}) and percent C_4 biomass contribution to soil CO_2 . Asterisk denotes nodules sampled at the LOM3 site, 2011-2b (see Extended Data Fig. 2a). **b**, Schematic box and whisker plots of f_{wc} from the LOM3 (3.3 Ma, this study) and Gona^{33,54,55} (Busidima Fm, 2.5–2.7 Ma) lithic sites and other East African hominin localities from 3.2–3.4 Ma^{54–61} relative to UNESCO structural categories of African vegetation^{32,52}. Grey box denotes 25th and 75th percentiles (interquartile range); whiskers represent observations within upper and lower

fences (1.5 \times interquartile range); black line shows mean value; grey line equals median value; black circles indicate mild outliers. **c**, Summary statistics of paleosol $\delta^{13}\text{C}_{\text{VPDB}}$ values and f_{wc} from LOM3 (3.3 Ma) and Gona^{33,54,55} (2.5–2.7 Ma) lithic sites and other East African hominin localities from 3.2–3.4 Ma^{54–61}. LOM3 $\delta^{13}\text{C}_{\text{VPDB}}$ values are significantly lower than those from the Busidima Formation at Gona (t -test, $P < 0.001$) and have a mean value that indicate 18% more woody canopy cover. When compared to paleosol $\delta^{13}\text{C}_{\text{VPDB}}$ values of the Koobi Fora, Nachukui, Chemeron, and Hadar formations from 3.2 to 3.4 Ma, LOM3 $\delta^{13}\text{C}_{\text{VPDB}}$ values are not significantly different (one-way ANOVA, $P > 0.05$).



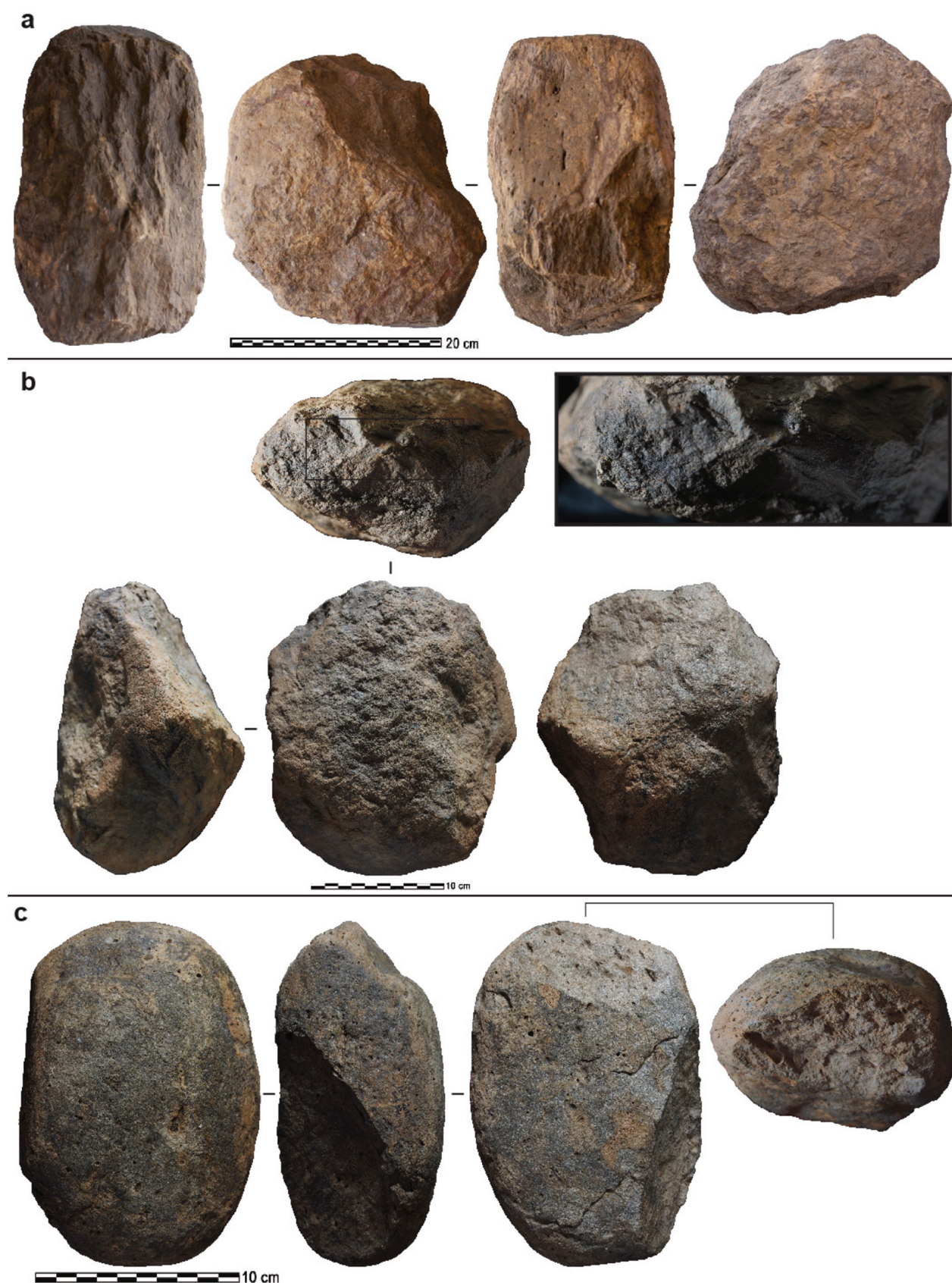
Extended Data Figure 5 | Gradual uncovering of core I16-3 from *in situ* pliocene sediment. **a**, Photograph showing square I16 at the beginning of excavation. Yellow line indicates north wall of square (July 16, 2011, 12.14 p.m.). **b**, Close-up of square I16 indicating complete burial of as-yet-uncovered artefact I16-3 (12.14 p.m.). **c**, Square I16 after excavation had begun and artefact I16-3 was initially exposed (2:11 p.m.). **d**, Close-up of artefact I16-3 after being

initially exposed (2.12 p.m.). **e**, Close-up of artefact I16-3 after further excavation (3.02 p.m.). **f**, Square I16 after further excavation (5.32 p.m.). **g**, Close-up of artefact I16-3 after further excavation (5.34 p.m.). **h**, Close-up of artefact I16-3 after being completely freed from the surrounding matrix and flipped over for inspection (5.36 p.m.). **i**, Close-up of impression from under artefact I16-3 (5.47 p.m.).



Extended Data Figure 6 | Photos of selected LOM3 artefacts compared with similar experimental cores. Together with the technological analysis of the archaeological material, our replication experiments suggest that the LOM3 knappers were using passive hammer technique, in which the core, usually held in both hands, is struck against a stationary object that serves as the percussor³⁴ (also referred to as on-anvil, block on block or *sur percuteur dormant*³⁵) and/or bipolar technique, in which the core is placed on an anvil and struck with a hammerstone³⁴. **a**, Unifacial passive hammer cores. Left is archaeological piece LOM3-2012 surf 106 (2.04 kg); right is experimental piece Expe 55 (3.40 kg) produced using the passive hammer technique. Selection of relatively flat blocks with natural obtuse angles. The flake removal process starts from a slightly prominent part of the block (white arrows show the direction of removals). The

removals tend to be invasive. The flaked surface forms a semi-abrupt angle with the platform surface. A slight rotation of the block ensures its semi-peripheral exploitation. **b**, Unifacial bipolar cores. Left are archaeological pieces LOM3-2012-H18-1 (left, 3.45 kg) and LOM3-2012 surf 64 (right, 2.58 kg); right are experimental pieces Expe 39 (left, 4.20 kg) and Expe 24 (right, 2.23 kg) produced using the bipolar technique. The block selected are thicker and more quadrangular in shape with natural angles $\approx 90^\circ$. Flakes are removed from a single secant platform (white arrows show the direction of removals). The flaked surface forms an abrupt angle with the other faces of the block. Impacts due to the contrecoups (white dots) are visible on the opposite edge from the platform.



Extended Data Figure 7 | Photographs of selected LOM3 artefacts.

a, Passive element/anvil (LOM3-2012 surf 50, 15 kg). Heavy sub-rectangular block displaying flat faces and therefore a natural morphology and weight which would enable stability. **b**, Hammerstone showing isolated impact points

(LOM3-2012 surf 33, 3.09 kg) and **c**, Hammerstone showing isolated impact points (LOM3-2012 surf 54, 1.63 kg), associated with a flake-like fracture on one end.

Extended Data Table 1 | Numerical data on the LOM3 lithic assemblage (2011, 2012).

a

ARTEFACTS		<i>in situ</i>	slope deposit	surface	total
core	passive hammer technique	1	1	33	35
	bipolar technique	2	3	26	31
	passive hammer and/or freehand technique	0	0	6	6
	bipolar and/or passive hammer technique	2	0	4	6
	of indeterminate technique	1	0	4	5
flake, whole or broken		10	3	22	35
percussor, whole, broken or potential		0	0	7	7
passive element/potential anvil		2	1	4	7
worked cobble		0	0	3	3
split cobble		0	0	2	2
fragment indet.		1	2	4	7
indet.		0	0	5	5
total		19	10	120	149

b

RAW MATERIAL	N	%
phonolite	51	34.23
trachy-phonolite	35	23.49
basalt	52	34.90
trachyte	2	1.34
vesicular basalt/conglomerate	3	2.01
indet.	6	4.03
total	149	100.00

a, Initial categorisation of the lithic components. b, Breakdown of lithic raw materials in the LOM3 assemblage.

Extended Data Table 2 | Comparison of whole flake and core dimensions between LOM3, early Oldowan sites and chimpanzee stone tool sites

Site	Age (Ma)	Ref.	N	Length				Width				Thickness			Geo. Mean	Mean Mass	
				Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min			Max
FLAKES																	
LOM3	3.3		26	120	48.8	19	205	110.1	40.7	19	185	43.9	23.4	6	90	59.9	842.4 (N=26)
OGS7	2.6	62	73	39.1*	14.3	13	80	37.1*	14.1	13	74	12.7*	5.07	3	26	14.10	18.9 (N=76) [‡]
EG10	2.6	62	114	37.38*	15.34	14	78	34.63*	13.74	14	78	13.18*	6.26	3	33	13.74	24.9 (N=72) [‡]
EG12	2.6	62	62	34.5*	12.84	15	66	35.55*	13.23	19	66	12.13*	5.76	4	30	13.23	21.5 (N=61) [‡]
AL894	2.36	63 [†]	1048	35.9*	23.63	6	134	25.07*	17.57	2	106	7.98*	6.4	1	45	17.1	
LA2C	2.34	16	500	38*	15	12	96	35*	14	7	128	11*	5	3	28	14.00	
Omo57	2.34	14 [†]	44	24.75*	10.546	10	58	20.36*	6.851	10	44	7.73*	4.008	1	18	6.85	
Omo123	2.34	14 [†]	110	20.8*	7.495	7	50	17.79*	6.485	6	38	5.9*	2.792	1	16	6.49	
DK	> 1.84	64	115	40.18*	14.803	18	111	37.41*	11.215	17	71	11.89*	5.404	4	29	11.22	
FLKZinj	1.76-1.84	64	125	36.78*	12.13	16	82	32.88*	11.59	4	76	11.51*	5.45	4	36	11.59	
Noulo [§]	.0043	40	5	35*	20.62	15	70	48*	27.06	15	90	11.6*	3.21	8	15	20.15	
CORES																	
LOM3	3.3		83	167	23.4	132	260	147.8	23.1	90	210	108.8	21.8	61	170	139	3096.4 (N=81)
OGS7	2.6	62	7	44.14*	13.68	28	67	59*	8.54	45	70	37*	8.2	22	49	45.85	78 [‡]
EG10	2.6	62	16	83.33*	10.34	69	105	60.9*	9.18	44	80	45.27*	12.36	30	69	61.25	232 [‡]
EG12	2.6	62	7	74.45*	8.72	58	93	59.73*	8.06	49	77	43.73*	77.4	25	53	57.94	194 (N=9) [‡]
AL894	2.36	63 [†]	38	75.01*	30.32	19.31	136.3	55.33*	22.54	12.21	94.9	35.87*	18.1	7.92	78.2	53.00	
LA2C	2.34	16	70	66*	18	39	123	52*	14	32	95	32*	12	12	78	47.9	
Omo 57	2.34	14 [†]	7	37.4*	8.81	25	52	28.8*	7.313	22	40	16.5*	4.721	11	24	26.10	
Omo 123	2.34	14 [†]	11	30.5*	12.193	17	56	22.27*	8.186	13	42	13.5*	4.569	9	24	20.93	
DK	> 1.84	64	69	67.93*	19.146	30	117	62.78*	17.992	25	100	48.25*	14.435	18	81	59.04	
FLK Zinj (lava only)	1.76-1.84	64	49	76.35*	12.57	53	95	78.85*	16.26	49	112	59*	12.3	37	87	70.82	

Dimensions are in mm, Mass in g. *Denotes significant difference with LOM3 (t-test, one-tailed, $P < 0.0001$). Given the small sample sizes and potential non-normal nature of stone tool measurements, a non-parametric test such as Mann-Whitney would be preferable, but this would require access to the raw measurement data from the other Oldowan sites, access to which is currently beyond the scope of this work. The Student's t-test is very robust, however, as deviations from normality do not affect it very much, and it is currently the only option when working with published data summaries.

[†]The summary data from this publication was not in the correct format for direct comparison with LOM3, so information for this table was provided directly by the author in the form of personal communication.

[‡]Data from ref. 17, hence differing sample sized from ref. 62.

[§]Dimensions of accidentally produced flakes from chimpanzee nut-cracking activity are included here for comparative purposes, although a direct technological comparison would be inappropriate as those pieces are not the result of intentional flake manufacture and do not bear the classic technological flake characteristics like those from LOM3 and early Oldowan sites.

Extended Data Table 3 | Comparison of anvils and percussors dimensions found at LOM3 site with anvils and percussors used by non-human primates in Bossou (wild chimpanzees, *Pan troglodytes verus* from ref. 41)

		Length		Width		Thickness		Mass	
Site	N	Mean	Std	Mean	Std	Mean	Std	Mean	Std
ANVILS									
LOM3	7	19.40	±6.26	15.44	±3.69	12.54	±3.17	6511.29	±4901.64
BOSSOU	32	16.10	±6.47	11.1*	±4.38	7.30*	±2.97	2200.00*†	±2210.00
PERCUSSORS									
LOM3	7	13.38	±3.67	10.75	±2.81	7.84	±2.33	1602.00	±1120.55
BOSSOU	35	12.00	±2.65	7.70*	±1.86	5.20*	±1.08	700.00*	±330.00

Dimensions are in cm, Mass in g. *Denotes significant difference with LOM3 (t-test, two-tailed, $P < 0.0199$). Given the small sample sizes and potential non-normal nature of stone tool measurements, a non-parametric test such as Mann–Whitney would be preferable but this would require access to the raw measurement data from ref. 41, access to which is currently beyond the scope of this work. The Student's *t*-test is very robust, however, as deviations from normality do not affect it very much and it is currently the only option when working with published data summaries.

† $N = 31$.

An alternative pluripotent state confers interspecies chimaeric competency

Jun Wu^{1*}, Daiji Okamura^{1*†}, Mo Li¹, Keiichiro Suzuki¹, Chongyuan Luo^{2,3}, Li Ma¹, Yupeng He³, Zhongwei Li¹, Chris Benner⁴, Isao Tamura¹, Marie N. Krause¹, Joseph R. Nery³, Tingting Du⁵, Zhuzhu Zhang³, Tomoaki Hishida¹, Yuta Takahashi^{1,6}, Emi Aizawa¹, Na Young Kim¹, Jeronimo Lajara⁷, Pedro Guillen^{7,8}, Josep M. Campistol⁹, Concepcion Rodriguez Esteban¹, Pablo J. Ross¹⁰, Alan Saghatelian¹¹, Bing Ren⁵, Joseph R. Ecker^{2,3} & Juan Carlos Izpisua Belmonte¹

Pluripotency, the ability to generate any cell type of the body, is an evanescent attribute of embryonic cells. Transitory pluripotent cells can be captured at different time points during embryogenesis and maintained as embryonic stem cells or epiblast stem cells in culture. Since ontogenesis is a dynamic process in both space and time, it seems counterintuitive that these two temporal states represent the full spectrum of organismal pluripotency. Here we show that by modulating culture parameters, a stem-cell type with unique spatial characteristics and distinct molecular and functional features, designated as region-selective pluripotent stem cells (rsPSCs), can be efficiently obtained from mouse embryos and primate pluripotent stem cells, including humans. The ease of culturing and editing the genome of human rsPSCs offers advantages for regenerative medicine applications. The unique ability of human rsPSCs to generate post-implantation interspecies chimaeric embryos may facilitate our understanding of early human development and evolution.

Two types of pluripotent stem cells (PSCs) have been captured from early mouse embryos. Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) of a pre-implantation blastocyst^{1,2} resemble naive epiblast³, and epiblast stem cells (EpiSCs) established from post-implantation epiblast are probably the *in vitro* counterparts of anterior primitive-streak cells^{4–6}. While both are pluripotent, they bear striking differences in molecular signature, signalling dependency, colony morphology, cloning efficiency, metabolic requirements and epigenetic features^{7,8}, which together with their ability to re-enter embryogenesis at different developmental time points (pre-implantation versus post-implantation, respectively) distinguish ESCs and EpiSCs as existing in two temporally distinct pluripotent states.

After embryo implantation, signals from regionalized extra-embryonic tissues guide pluripotent epiblast cells through dynamic changes to initiate the embryonic body plan that accommodates the diversified developmental fates that ensue upon gastrulation⁹. Heterotopic grafting experiments indicate that epiblast cells, regardless of their regional origins, can adopt the developmental fate characteristic of the cell population at the site of transplantation, illustrating their highly plastic nature¹⁰. Nonetheless, it is conceivable that epiblasts are subjected to regional influences and bear a multitude of pluripotent states with distinguishable molecular and functional signatures¹¹. To date it remained unknown whether PSCs with distinct spatial identities could be stabilized in culture. By carefully examining the cellular response of the epiblast to different *ex vivo* environmental stimuli, we have isolated, with high efficiency, a stable primed pluripotent cell type from both pre- and post-implantation epiblasts that differs from EpiSCs in cloning efficiency, cell growth kinetics, transcriptomic, epigenomic and metabolic profiles. Notably, the newly identified PSCs selectively

colonize the posterior region of post-implantation embryos and allow for efficient generation of *ex vivo* intra- and interspecies chimaeric embryos. Our study not only uncovers a novel spatially defined pluripotent cell type, but also opens up a new avenue for comparing early developmental programs across species.

Optimizing epiblast culture parameters

FGF2/Activin-A (F/A) signalling supports the derivation of EpiSCs^{4,5,12}. While deriving EpiSCs using a F/A-based medium¹³, we observed cellular differentiation started around day 3 and by day 4 only a few undifferentiated epiblast cells remained (Fig. 1a, b and Extended Data Fig. 1a, b). This suggested to us that the pluripotent states of most of the cells present across the *in vivo* epiblast could not be maintained by F/A signalling. The canonical Wnt signalling pathway also has an important role in EpiSC self-renewal^{14–17}. We tested the effect of a Wnt inhibitor IWR1 on epiblast explants. Isolated E5.75 epiblasts were cultured in a serum-free N2B27 medium¹⁸ on mitotically inactivated mouse embryonic fibroblasts (MEFs) supplemented with IWR1 (N2B27^{R1}) (Fig. 1a, b). After 4 days in culture, we found the number of SSEA-1⁺/OCT4⁺ cells dramatically increased in N2B27^{R1} compared to F/A-based medium (Extended Data Fig. 1b). However, a significant fraction of SSEA-1⁺/OCT4⁺ cells was still detected. Next we tested the combination of either Activin-A/IWR1 (N2B27^{A/R1}) or FGF2/IWR1 (N2B27^{F/R1}). Notably, while a comparable level of differentiation was observed in N2B27^{A/R1} versus N2B27^{R1}, day 4 epiblast outgrowths in N2B27^{F/R1} showed homogeneous morphology and little-to-no differentiation (Fig. 1b, c and Extended Data Fig. 1b, c). Mechanistically, the combination of the serum-free N2B27 medium, IWR1 and FGF2 suppressed lineage

¹The Salk Institute for Biological Studies, Gene Expression Laboratory, La Jolla, California 92037, USA. ²Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California 92037, USA. ³The Salk Institute for Biological Studies, Genomic Analysis Laboratory, La Jolla, California 92037, USA. ⁴The Salk Institute for Biological Studies, Integrated Genomics, La Jolla, California 92037, USA. ⁵Ludwig Institute for Cancer Research, University of California, San Diego School of Medicine, Department of Cellular and Molecular Medicine, 9500 Gilman Drive, La Jolla, California 92093-0653, USA. ⁶Life Science Center, Tsukuba Advanced Research Alliance, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8577, Japan. ⁷Grado en Medicina, Universidad Católica, San Antonio de Murcia, Campus de los Jerónimos, 135, Guadalupe 30107, Spain. ⁸Fundacion Pedro Guillen, Clínica Centro, Avenida Ventisquero de la Condesa, 42, 28035 Madrid, Spain. ⁹Hospital Clinic of Barcelona, Carrer Villarroel, 170, 08036 Barcelona, Spain. ¹⁰University of California, Davis, Davis, California 95616, USA. ¹¹The Salk Institute for Biological Studies, Peptide Biology Laboratory, La Jolla, California 92037, USA. [†]Present address: Department of Advanced Bioscience, Graduate School of Agriculture, Kinki University, 3327-204 Nakamachi, Nara 631-8505, Japan.

*These authors contributed equally to this work.

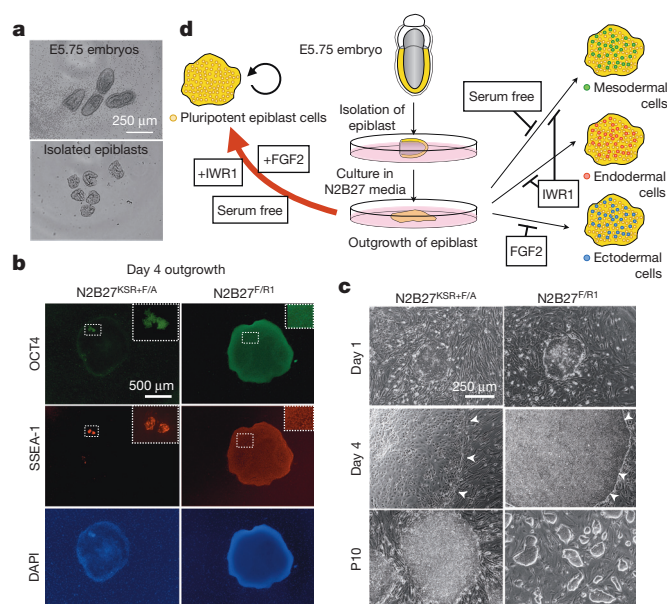


Figure 1 | The effects of culture parameters on epiblast explants. **a**, Freshly collected E5.75 mouse embryos (top) and isolated E5.75 epiblasts (bottom). **b**, Immunostaining of day 4 epiblast outgrowths. KSR, KnockOut serum replacement. Green, OCT4; red, SSEA-1; blue, DAPI. Insets, higher-magnification images. **c**, Morphologies of epiblast outgrowths at day 1 and day 4 in EpiSC and rsEpiSC derivation medium. Bottom, EpiSCs and rsEpiSCs at passage 10 (P10). Arrowheads, edges of the epiblast outgrowths. **d**, Schematic model summarizing the effects of various culture parameters on epiblast explants.

differentiation and arrested the majority of, if not all, epiblast cells in a proliferative state, with homogenous expression of the pluripotency markers OCT4 and SSEA-1 (Fig. 1d and Extended Data Fig. 2a, b).

A spatially defined pluripotent state

Upon passaging with collagenase type IV, traditionally used for EpiSCs, we could derive stable cell lines under N2B27^{F/R1}, referred to as EpiSCs^{F/R1}. Surprisingly, EpiSCs^{F/R1} could also be efficiently derived after trypsin disaggregation of day 4 epiblast outgrowths (Extended Data Fig. 3a). In addition to IWR1, we could also obtain stable cell lines using other Wnt inhibitors XAV939 and IWP2 (Extended Data Fig. 3i, j). In our experiments, the derivation success rate with F/A culture is around 33%, similar to a recent report⁶. In contrast, we could readily obtain stable EpiSCs^{F/R1} from different genetic backgrounds and different developmental stages of post-implantation as well as pre-implantation epiblasts, even after the first passage with a derivation success rate of 100% (Extended Data Fig. 3b, c, e), a feat not possible with F/A culture^{5,19}. Moreover, N2B27^{F/R1} equally supported the derivation of EpiSCs^{F/R1} from four micro-dissected quadrants (anterior-proximal, anterior-distal, posterior-proximal and posterior-distal) of E6.5 epiblasts with a perfect success rate (Extended Data Fig. 3f, g). This uniform response suggests that N2B27^{F/R1} captures a pluripotent state accessible to all *in vivo* pluripotent epiblast cells of diverse spatiotemporal origins.

EpiSCs^{F/R1} could be maintained long term in culture while displaying homogenous morphology and a normal karyotype (Fig. 1c and Extended Data Fig. 4f). EpiSCs^{F/R1} derived from different developmental stages and different regions exhibited gene expression patterns characteristic of the primed-state PSCs (Extended Data Figs 3d, h and 4a). EpiSCs^{F/R1} expressed standard pluripotent protein markers and possessed weak alkaline phosphatase activity (Extended Data Fig. 4b, c, g). Epigenetically, EpiSCs^{F/R1} exhibited X-chromosome inactivation in female cells (Fig. 2a and Extended Data Fig. 4b), demethylated *Oct4* (also known as *Pou5f1*) promoter and fully methylated *Stella* (also known as *Dppa3*) and *Dppa5* promoters (Extended Data

Fig. 4d). We found that self-renewal of EpiSCs^{F/R1} was dependent on balanced signalling of FGF2 and IWR1 (Extended Data Fig. 5a–c). Pluripotency of EpiSCs^{F/R1} was demonstrated with *in vivo* teratoma assays and we observed that the sizes of teratomas derived from EpiSCs^{F/R1} were smaller than those derived from EpiSCs (Extended Data Fig. 4i–k). Collectively, these results indicate that EpiSCs^{F/R1} bear properties characteristic of the primed pluripotent state.

Low clonogenicity is regarded as one of the prominent features of primed EpiSCs. In contrast, EpiSCs^{F/R1} showed high cloning efficiency (~34.2%) at a comparable level to ESCs (~41.3%) and much higher than EpiSCs (~1.1%) (Fig. 2b and Extended Data Fig. 4e)^{15,20}. Other notable differences between EpiSCs^{F/R1} and EpiSCs are proliferation rate and doubling time. EpiSCs^{F/R1} proliferate at a much faster pace than EpiSCs (Fig. 2c). The doubling time of EpiSCs^{F/R1} (8–10 h) is markedly shorter than that of EpiSCs (14–16 h) and this is probably due to a shortened G1 phase (Extended Data Fig. 4h). Metabolically, we found EpiSCs^{F/R1} were more dependent on glycolysis and less on mitochondrial respiration than already highly glycolytic EpiSCs⁸ (Extended Data Fig. 7a–g), which might support their fast growth rate. The high cloning efficiency, rapid proliferation rate and a more glycolytic metabolic profile suggest that EpiSCs^{F/R1} exist in a primed pluripotent state distinct from that of the conventional EpiSCs.

Naive mouse ESCs (mESCs) differ from primed EpiSCs in their ability to generate chimaeras following blastocyst injection. Post-implantation embryos, however, constitute a non-permissive environment for ICM-derived mESCs, and hence, grafted mESCs proliferate poorly²¹. To phenotypically evaluate the pluripotent state of EpiSCs^{F/R1}, we first performed blastocyst injections. We did not observe any chimaera contribution from EpiSCs^{F/R1}, further supporting the notion that EpiSCs^{F/R1} are in a primed pluripotent state. Although incapable of colonizing pre-implantation ICMs, EpiSCs could readily incorporate and generate chimaeras when grafted into post-implantation epiblasts followed by *in vitro* embryo culture²¹. To functionally define EpiSCs^{F/R1} we grafted Kusabira-Orange-labelled EpiSCs^{F/R1} into different regions of isolated, non-intact E7.5 embryos (anterior, distal and posterior) (Fig. 2d). Unlike conventional EpiSCs, which incorporated efficiently in the distal (10/12) and posterior (8/10) regions and to a lesser extent the anterior region (6/10), an observation consistent with previous reports^{6,15,21}, EpiSCs^{F/R1} only integrated efficiently in the posterior epiblast (22/25). They poorly integrated in the anterior region (3/12) and not at all in the distal region (0/12). More importantly, after 36 h of embryo culture only EpiSCs^{F/R1} grafted to the posterior region could disperse from graft sites, proliferate and differentiate into the three germ layers in chimaeric embryos (Fig. 2e–g and Extended Data Fig. 4m). These results not only confirmed the pluripotency of EpiSCs^{F/R1}, but also revealed a preferential affinity and high compatibility between EpiSCs^{F/R1} and the posterior E7.5 epiblast. Based on this unique embryo grafting property, we named EpiSCs^{F/R1} as region-selective EpiSCs, or rsEpiSCs. Divergent grafting outcomes indicate that rsEpiSCs represent a class of primed-state PSCs with a new spatial identity distinct from conventional EpiSCs.

Multiple omics comparisons of PSCs

We compared the transcriptomes of mESCs, EpiSCs, rsEpiSCs and *in vivo* isolated epiblasts using microarrays. Principle component analysis (PCA) and unsupervised hierarchical clustering analysis showed that rsEpiSCs clustered tightly as a group separated from both mESCs and EpiSCs, indicating that rsEpiSCs acquired a distinct global transcriptome profile (Fig. 3a and Extended Data Fig. 6a). Comparative analysis of RNA-seq data identified 2,245 genes differentially expressed between rsEpiSCs and EpiSCs using a fourfold cut-off (FDR < 0.05, Extended Data Fig. 6b and Supplementary Table 1). Notable gene ontology (GO) terms enriched in genes upregulated in rsEpiSCs were related to neuron differentiation and development (Extended Data Fig. 6c). FACS analyses of teratomas, however, revealed no obvious lineage biases between EpiSCs and rsEpiSCs

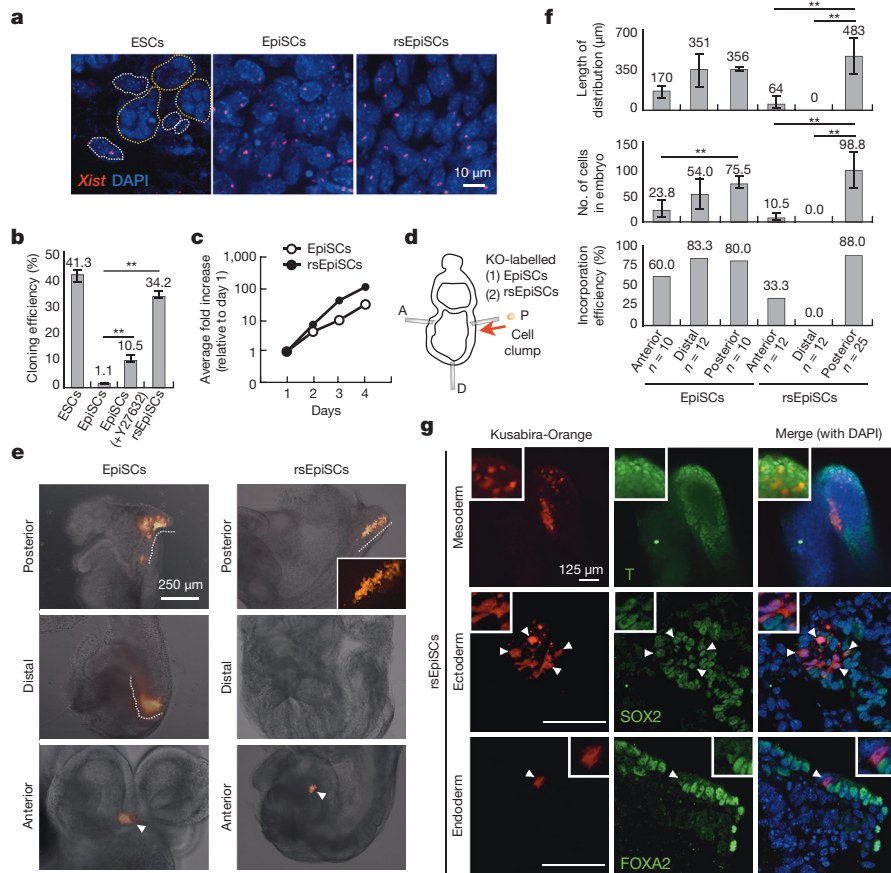


Figure 2 | Characterization of rsEpiSCs. **a**, *Xist* RNA FISH signals in female mouse ESCs, EpiSCs and rsEpiSCs. Dotted circles: yellow, mESCs; white, MEFs. **b**, Cloning efficiencies of mESCs, EpiSCs and rsEpiSCs. **c**, Growth curve of EpiSCs and rsEpiSCs. **d**, Schematic representation of epiblast grafting experiments (please refer to Supplementary Fig. 1 for details). KO, Kusabira-Orange. A, anterior; P, posterior; D, distal. **e**, Representative images showing outcomes of grafted Kusabira-Orange-labelled EpiSCs and rsEpiSCs to

different epiblast regions. Dashed lines, dispersed cells. Arrowheads, cell clumps. **f**, Distinct outcomes for EpiSCs and rsEpiSCs grafted to different epiblast regions. **g**, Whole-mount immunostaining of posterior grafted rsEpiSCs. Blue, DAPI. Arrowheads, SOX2 or FOXA2-positive derivatives of grafted cells. Top-middle, T (brachyury)-positive derivatives of grafted cells. Insets, higher-magnification images. Error bars, s.d.; *t*-test, $^{**}P < 0.01$ (**b**, $n = 6$; **c**, $n = 2$; **f**, n indicated on the graph, independent experiments)

(Extended Data Fig. 4l). Other prevalent GO terms were associated with cell motion, extracellular matrix and plasma membrane. This is reminiscent of the epithelial to mesenchymal transition (EMT) that takes place during gastrulation. We further explored this possibility by analysing the levels of E-CADHERIN, CLAUDIN-3 and SNAIL, proteins integral to the EMT process. Notably, we detected reduced expressions of both E-CADHERIN and CLAUDIN-3 and a concomitant increase in SNAIL expression, suggesting that the EMT process might have been initiated in rsEpiSCs (Fig. 4b). It is interesting to note that, consistent with the metabolic flux analysis, gene expression levels of enzymes and complexes involved in glycolysis were higher in rsEpiSCs versus EpiSCs, and those related to mitochondrial function were lower in rsEpiSCs versus EpiSCs (Extended Data Fig. 7h).

ChIP-seq analysis revealed that the global histone 3 lysine 4 trimethylation (H3K4me3) distribution pattern was similar between EpiSCs and rsEpiSCs while a significant increase of H3K27me3 levels was detected at the transcription start site of polycomb group target genes in rsEpiSCs (Fig. 3b, Extended Data Fig. 6d, <http://neomorph.salk.edu/rsEpiSC/browser.html>). The genomes of both EpiSCs and rsEpiSCs are highly methylated in CG contexts ($\sim 87\%$, Extended Data Fig. 6f). The genome-wide discovery of differentially methylated regions (DMRs) identified 1,336 DMRs between EpiSCs and rsEpiSCs with the vast majority (88.3%) showing hyper-methylation in rsEpiSCs. Of these rsEpiSCs hyper-DMRs, 53.8% (635/1,180) were located within 2.5 kilobases of the transcription start site (Extended Data Fig. 6f). rsEpiSCs hyper-DMRs were strongly enriched at

CpG-island-containing promoters ($P = 2.14 \times 10^{-88}$). A subset of genes (cluster 3) associated with complete promoter methylation and concomitant loss of H3K4me3 in rsEpiSCs was accompanied by reduced expression levels (48% of genes in cluster 3, $P = 1.6 \times 10^{-8}$, Fig. 3c). Consistent with the selective engraftment of rsEpiSCs to the posterior epiblast, genes associated with rsEpiSCs hyper-DMRs were enriched for GO terms such as regionalization and anterior/posterior pattern specification (Extended Data Fig. 6h). In both EpiSCs and rsEpiSCs, we observed a positive correlation between gene expression and non-CG methylation levels in the gene body, characteristic of pluripotent cells²² (Extended Data Fig. 6g). Consistent with transcriptomic analysis, genes related to cell membrane and neuronal lineage exhibited distinct patterns while pluripotency-related genes remained comparable in DNA methylation and/or histone methylation levels between EpiSCs and rsEpiSCs (Extended Data Fig. 6e).

Untargeted metabolomics and lipidomics analysis quantified differences in hydrophilic and hydrophobic metabolites, respectively, between EpiSCs and rsEpiSCs. Several hundred metabolites were identified as being different between the two cell lines using XCMS online²³ and the METLIN database²⁴ (Extended Data Fig. 7i, j and Supplementary Tables 2–7). The data show broad changes in metabolite levels including increased tricarboxylic acid cycle intermediates and lower levels of lipids in EpiSCs, indicating increased energy utilization in these processes. The increased tricarboxylic acid cycle usage is in line with a higher mitochondria respiration rate in EpiSCs than

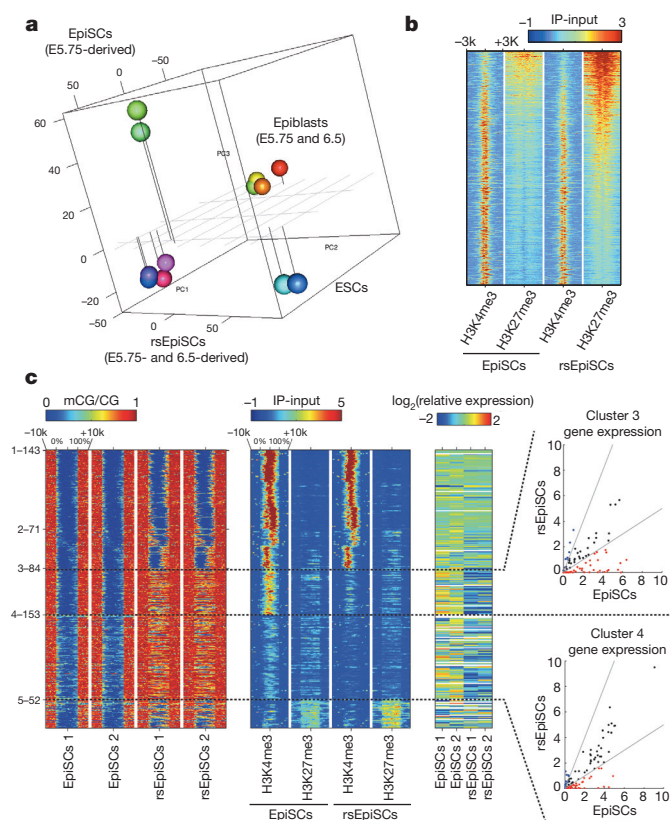


Figure 3 | Global transcriptomic and epigenomic analysis. **a**, PCA of microarray data generated from EpiSCs, rsEpiSCs and isolated epiblasts. **b**, H3K4me3 and H3K27me3 ChIP-seq signals at Polycomb target genes in EpiSCs and rsEpiSCs. **c**, Clustering of unmethylated regions associated with promoters that overlap with rsEpiSCs hyper-DMRs. CG methylation (mCG), H3K4me3 and H3K27me3 levels were plotted for scaled unmethylated regions and ± 10 kb regions surrounding the unmethylated regions. Dashed lines in the scatter plot indicate a twofold up (blue) or down (red) difference in RNA level in rsEpiSCs compared to EpiSCs. **a**, **c**, $n = 2$; **b**, $n = 1$, biological replicates.

rsEpiSCs. Also, rsEpiSCs contained less glucose and more glucose-6-phosphate compared to EpiSCs, which correlates with the higher glycolytic activity in rsEpiSCs as determined by medium acidification measurements.

Collectively these multi-omic results distinguish rsEpiSCs from EpiSCs at the transcriptomic, epigenomic and metabolic levels, further underlining their contrasting attributes despite both existing in the primed pluripotent state.

In vivo relevance of rsEpiSCs

To investigate which region(s) of the *in vivo* epiblast rsEpiSCs may resemble, we performed RNA-seq on four dissected regions of the late E6.5 epiblast (anterior-proximal, anterior-distal, posterior-proximal and posterior-distal) and compared differentially expressed genes among *in vivo* samples with rsEpiSCs. Spearman's rank correlation revealed that *in vivo* posterior-proximal epiblast had a higher correlation with rsEpiSCs than other epiblast quadrants, suggesting that rsEpiSCs might have acquired cellular properties characteristic of the posterior-proximal epiblast (Fig. 4a).

To reveal the temporal identity of rsEpiSCs, we first tested the ability of rsEpiSCs to be induced into primordial germ cells (PGCs). Epiblasts acquire the ability to adopt a PGC fate in response to BMP4 between developmental stages E5.5 and E6.5^{25,26}. We did not observe any PGC induction in rsEpiSCs following an established protocol¹³ (Extended Data Fig. 6l), suggesting that they do not resemble *in vivo* epiblasts during this period. To help dissect which developmental stage rsEpiSCs most closely relate to, we next compared the transcriptomes

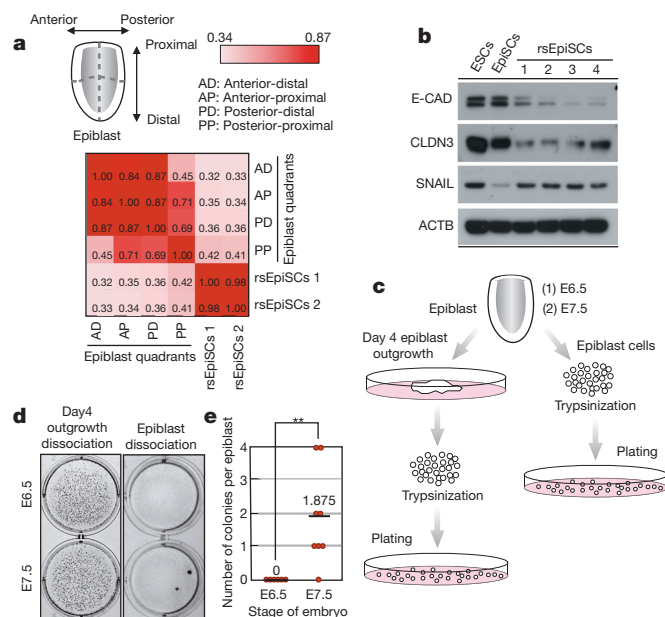


Figure 4 | In vivo relevance of rsEpiSCs. **a**, Transcriptomically rsEpiSCs correlate better with the posterior-proximal than with other regions of epiblast. **b**, Protein expression of EMT markers, ACTB served as the loading control. **c**, Schematic workflow for the clonal derivation experiments. **d**, Representative images showing OCT4-positive colonies after trypsinizing day 4 epiblast outgrowths (left) and isolated epiblast (right) from E6.5 and E7.5 mouse embryos. **e**, Number of OCT4-positive colonies per embryo from clonal derivation of rsEpiSCs. Error bars indicate s.d. t -test, $**P < 0.01$. (E6.5, $n = 6$; E7.5, $n = 8$, independent experiments). For the full scan associated with **b**, refer to Supplementary Information.

of rsEpiSCs to those of epiblasts isolated from different embryological timelines from a published data set⁶. PCA and hierarchical clustering analysis revealed that gene expression profiles of rsEpiSCs are similar to epiblast from late-streak/no-bud-stage embryos²⁷ (Extended Data Fig. 6i, j). This prompted us to investigate whether rsEpiSCs acquired primitive streak cellular properties by comparing rsEpiSCs and *in vivo* late-streak epiblast transcriptomes for an annotated list of primitive marker genes^{6,28,29}. We found almost all primitive streak-related genes were expressed at significantly lower levels in rsEpiSCs compared to *in vivo* late-streak epiblast, indicating rsEpiSCs had not acquired the molecular properties of primitive streak cells (Extended Data Fig. 6k).

To functionally time rsEpiSCs, we focused on their unusually high cloning efficiency and tested clonal derivation of rsEpiSCs by directly trypsinizing isolated epiblasts. When starting with E7.5 epiblasts, a few OCT4-positive colonies appeared after 7 day's culture. In contrast, we did not observe any colony forming with E6.5 epiblasts (Fig. 4c–e). These results indicate that some *in vivo* E7.5 epiblast cells acquired resistance to apoptosis after single-cell enzymatic dissociation, a property shared by rsEpiSCs. This may be attributed to the EMT occurring during gastrulation when posterior-proximal epithelial epiblast cells delaminate and ingress through the primitive streak to form mesoendoderm³⁰. Taken together, these findings suggest that rsEpiSCs may represent a subpopulation of cells from the late-streak/no-bud-stage epiblast undergoing EMT before their lineage commitment.

Primate region-selective PSCs

Notwithstanding their blastocyst origin, human ESCs exist in a primed pluripotent state similar to mouse EpiSCs, a state that suffers from several practical disadvantages, including low derivation and cloning efficiency. When human ESCs were cultured in F/R1 conditions we observed long-term self-renewal and karyotypic stability (Fig. 5a and Extended Data Fig. 8e). F/R1 human ESCs (designated as human region-selective ESCs, or human rsESCs) expressed

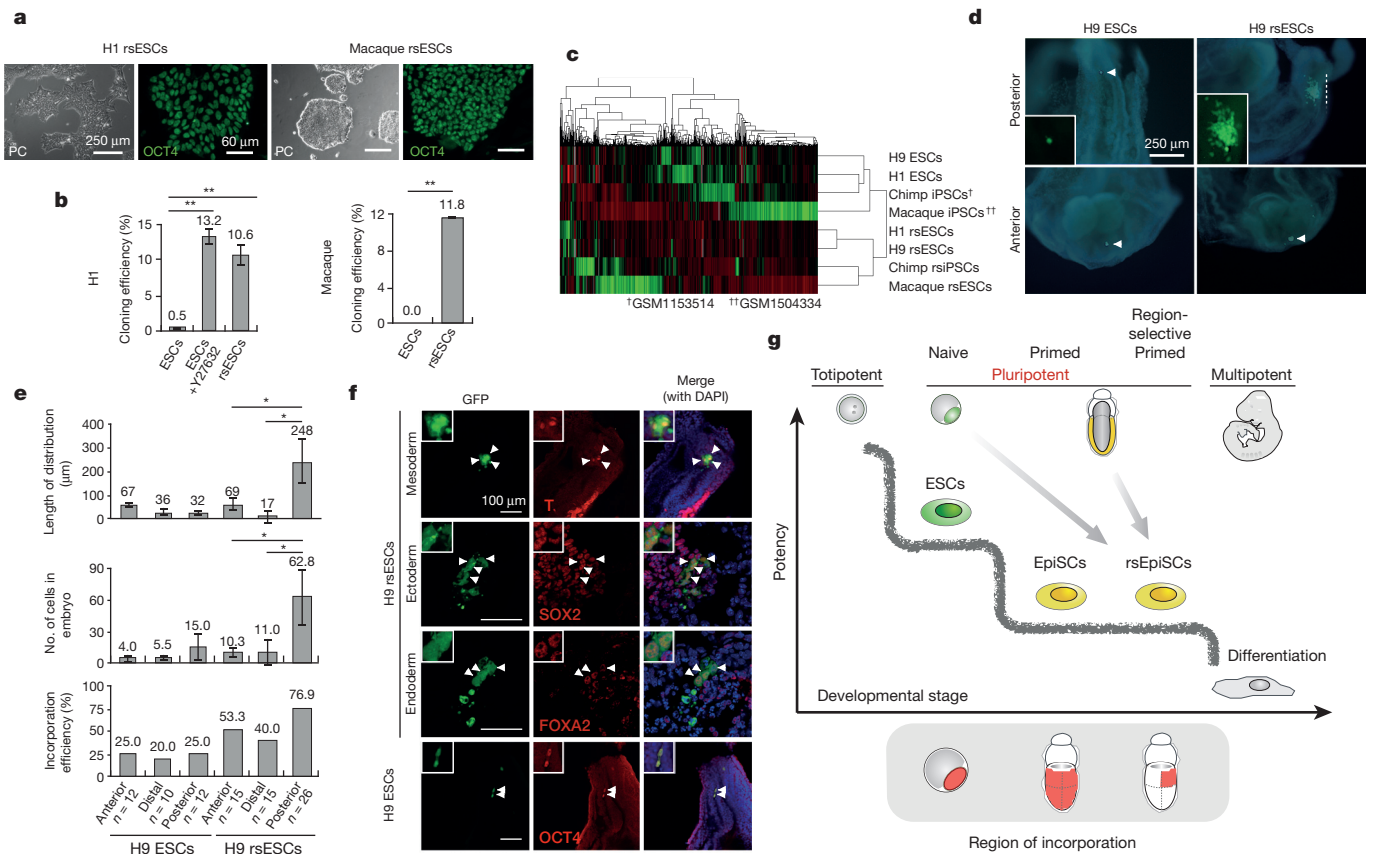


Figure 5 | Primate region specific PSCs. **a**, Morphology and OCT4 immunostaining of human and rhesus macaque rESCs. **b**, Cloning efficiency of human and macaque ESCs versus rESCs. **c**, Unbiased cross-species hierarchical clustering of transcriptomes of primate PSCs and rPSCs. **d**, Representative images showing outcomes of grafted GFP-labelled H9 ESCs and H9 rsESCs to different epiblast regions of non-intact and non-viable isolated E7.5 mouse embryos. Arrowheads, cell clumps. Dashed line, dispersed cells. Blue, DAPI. **e**, Outcomes of grafting H9 and H9 rsESCs to different mouse

epiblast regions of non-intact and non-viable E7.5 mouse embryos. **f**, Immunostaining of posterior-grafted GFP-labelled H9 ESCs and H9 rsESCs. Blue, DAPI. Insets, high magnification views. Arrowheads indicate T (brachyury)-, SOX2-, FOXA2- or OCT4-positive derivatives of grafted cells. **g**, Schematic representation of distinct PSCs captured from early embryos and their specific timings and regions to re-enter embryogenesis. Error bars, s.d.; *t*-test, $^{**}P < 0.01$, $^{*}P < 0.05$. (**b**, $n = 6$; **e**, n indicated on the graph; independent experiments).

standard pluripotency markers (Fig. 5a and Extended Data Fig. 8a, c, m), harboured a reduced population of cells retained at the G1 phase (Extended Data Fig. 8l) and generated teratomas in NOD/SCID mice comprising the three germ lineages (Extended Data Fig. 8d). Similar to mouse rsEpiSCs, the cloning efficiency of human rsESCs was significantly improved (Fig. 5b and Extended Data Fig. 8b), which greatly facilitated genome editing at a comparable level to Y27632-treated conventional human ESCs (Extended Data Fig. 9a–d). F/R1 culture also supported generation of human induced pluripotent stem cells (iPSCs). Notably, compared to F/A culture, putative iPSC-like colonies in F/R1 culture were more homogenous, bigger in size and contained less colonies with partial/absent alkaline phosphatase activity (Extended Data Fig. 8g–i).

F/R1 culture also supported long-term culture of non-human primate (NHP) PSCs including rhesus macaque PSCs³¹ and chimpanzee iPSCs³² (Fig. 5a and Extended Data Fig. 10a, b). The cloning efficiency of NHP rsESCs was also improved (Fig. 5b). Hierarchical clustering of the transcriptome analysed using RNA-seq showed that primate rsPSCs were clustered together in a group distinct from conventional primate PSCs in F/A culture (Fig. 5c).

To functionally test whether primate rsPSCs acquired phenotypic properties characteristic of mouse rsEpiSCs, we grafted green fluorescent protein (GFP)-labelled human H9 ESCs or H9 rsESCs into the anterior, distal and posterior regions of epiblasts of isolated non-intact E7.5 mouse embryos (see Methods). Conventional H9 ESCs could not efficiently integrate and proliferate inside mouse epiblasts regardless of

their grafting sites. In rare cases when few H9 ESCs were detected, they remained undifferentiated, as indicated by positive staining for OCT4 (Fig. 5d–f). This indicates that unlike mouse EpiSCs, human H9 ESCs are incompatible with mouse post-implantation epiblasts. In contrast, H9 rsESCs efficiently integrated, proliferated and differentiated into all three germ layers when grafted in the posterior region, remained undifferentiated as clusters in the anterior region and showed no incorporation in the distal region of the E7.5 mouse epiblasts, consistent with mouse rsEpiSCs (Fig. 5d–f). In addition to human cells, we observed similar engrafting patterns with GFP-labelled rhesus macaque rsESCs (Extended Data Fig. 10d–f), indicating that F/R1 culture endowed similar phenotypic features to NHP ESCs.

Pluripotency, an evanescent feature of early embryonic development, permeates epiblast cells of diverse spatiotemporal origins³³. Nevertheless, distinct pluripotent states have been described and distinguished along the embryological timeline without regard to spatial attributes. Although post-implantation pluripotent epiblast cells are not irreversibly committed, they are spatially polarized under the influence of local milieu provided by surrounding extra-embryonic tissues³⁴. The regional properties of post-implantation epiblast cells have largely been overlooked in EpiSCs. By refining culture parameters, we have captured cells of spatially defined pluripotency that harbour molecular, epigenetic and metabolic signatures that differ from conventional EpiSCs. Our study sets a precedent for exploring other spatially distinct pluripotent states in the early embryo.

Compared to naive cells, primed PSCs exist in a more developmentally advanced state and are poised for rapid and efficient differentiation. This advantage, however, is overshadowed by the heterogeneity and poor cloning efficiency associated with the conventional F/A culture. F/R1 culture boosts cloning efficiency, simplifies routine cultivation, improves the quality of iPSC generations, and facilitates genome editing in primed human PSCs. These features are attractive for a myriad of applications such as synchronized and efficient differentiation, large-scale cell production, and gene-correction for disease modelling and therapeutic purposes. In addition, the ability of human rESCs to differentiate into derivatives of all three embryonic germ layers in chimaeric embryos provides us with a novel platform to study early human developmental events that are otherwise difficult to investigate, such as gastrulation and early lineage commitment, presumed, but not yet fully proven, to be conserved between human and other animal model organisms. The capture and engineering of PSCs with distinct features^{35,36} may enrich our fundamental understanding of pluripotency in mammalian development and evolution, as well as expand the repertoire of cellular tools that can be harnessed for regenerative medicine applications.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 August 2014; accepted 18 March 2015.

Published online 6 May 2015.

- Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
- Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638 (1981).
- Martello, G. & Smith, A. The nature of embryonic stem cells. *Annu. Rev. Cell Dev. Biol.* **30**, 647–675 (2014).
- Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
- Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
- Kojima, Y. *et al.* The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* **14**, 107–120 (2014).
- Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
- Zhou, W. *et al.* HIF1 α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J.* **31**, 2103–2116 (2012).
- Tam, P. P. L. & Loebe, D. A. F. Gene function in mouse embryogenesis: get set for gastrulation. *Nature Rev. Genet.* **8**, 368–381 (2007).
- Tam, P. P. & Zhou, S. X. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev. Biol.* **178**, 124–132 (1996).
- Chenoweth, J. G., McKay, R. D. G. & Tesar, P. J. Epiblast stem cells contribute new insight into pluripotency and gastrulation. *Dev. Growth Differ.* **52**, 293–301 (2010).
- Greber, B. *et al.* Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* **6**, 215–226 (2010).
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).
- Tsakiridis, A. *et al.* Distinct Wnt-driven primitive streak-like populations reflect *in vivo* lineage precursors. *Development* **141**, 1209–1221 (2014).
- Sumi, T., Oki, S., Kitajima, K. & Meno, C. Epiblast ground state is controlled by canonical Wnt/ β -catenin signaling in the postimplantation mouse embryo and epiblast stem cells. *PLoS ONE* **8**, e63378 (2013).
- Berge, D. & *et al.* Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nature Cell Biol.* **13**, 1–8 (2011).
- Kim, H. *et al.* Modulation of β -catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nat. Commun.* **4**, 2403 (2013).
- Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
- Najm, F. J. *et al.* Isolation of epiblast stem cells from preimplantation mouse embryos. *Cell Stem Cell* **8**, 318–325 (2011).
- Watanabe, K. *et al.* A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature Biotechnol.* **25**, 681–686 (2007).
- Huang, Y., Osorno, R., Tsakiridis, A. & Wilson, V. *In vivo* differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Rep.* **2**, 1571–1578 (2012).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Tautenhahn, R., Patti, G. J., Rinehart, D. & Siuzdak, G. XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal. Chem.* **84**, 5035–5039 (2012).
- Tautenhahn, R. *et al.* An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnol.* **30**, 826–828 (2012).
- Okamura, D., Kimura, T., Nakano, T. & Matsui, Y. Cadherin-mediated cell interaction regulates germ cell determination in mice. *Development* **130**, 6423–6430 (2003).
- Okamura, D., Hayashi, K. & Matsui, Y. Mouse epiblasts change responsiveness to BMP4 signal required for PGC formation through functions of extraembryonic ectoderm. *Mol. Reprod. Dev.* **70**, 20–29 (2005).
- Downs, K. M. & Davies, T. Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* **118**, 1255–1266 (1993).
- Alev, C. *et al.* Transcriptomic landscape of the primitive streak. *Development* **137**, 2863–2874 (2010).
- Fossat, N., Pfister, S. & Tam, P. P. L. A transcriptome landscape of mouse embryogenesis. *Dev. Cell* **13**, 761–762 (2007).
- Acloque, H., Adams, M. S., Fishwick, K., Bronner-Fraser, M. & Nieto, M. A. Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *J. Clin. Invest.* **119**, 1438–1449 (2009).
- Mitalipov, S. *et al.* Isolation and characterization of novel rhesus monkey embryonic stem cell lines. *Stem Cells* **24**, 2177–2186 (2006).
- Marchetto, M. C. N. *et al.* Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* **503**, 525–529 (2013).
- Pera, M. F. In search of naivety. *Cell Stem Cell* **15**, 543–545 (2014).
- Beddington, R. S. & Robertson, E. J. Axis development and early asymmetry in mammals. *Cell* **96**, 195–209 (1999).
- Wu, J. & Belmonte, J. C. I. Stem cells: a designer's guide to pluripotency. *Nature* **516**, 172–173 (2014).
- Tong, P. D. *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* **516**, 192–197 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank S. Mitalipov and J. Thomson for providing rhesus ESCs and iPSCs, respectively; F. Gage for providing chimpanzee iPSCs; K. Zhang for assistance with cell line derivation; M. Ku of the H. A. and Mary K. Chapman Charitable Foundations Genomic Sequencing Core for performing RNA-seq and mouse ChIP-seq experiments; M. Chang of the Integrative Genomic and Bioinformatics Core for bioinformatics analysis; W. T. Berggren and the staff of the Salk STEM Core for preparation of custom-mTeSR1 base medium and supply of validated stem culture materials; G. Pao and K. Hasegawa for discussions; Y. Dayn from transgenic core facility and J. Luo for blastocyst injections; Y. Tsunekawa for providing the mutant eGFP human ESCs reporter line; E. O'Connor and K. Marquez of Human Embryonic Stem Cell Core Facility of Sanford Consortium for Regenerative Medicine for FACS sorting; R. H. Benitez, A. Goebel, R. D. Soligalia for assistance with genome editing; M. F. Pera for critical reading of the manuscript; and M. Schwarz, and P. Schwarz for administrative help. M.L. and K.S. are supported by a California Institute for Regenerative Medicine Training Grant. We thank J. L. Mendoza for his support on this project. This work was funded in part by UCAM (mouse studies). J.R.E. is an Investigator of the Howard Hughes Medical Institute. P.G. was supported by Fundacion Pedro Guillen. Work in the laboratory of J.C.I.B. was supported by G. Harold and Leila Y. Mathers Charitable Foundation, The Leona M. and Harry B. Helmsley Charitable Trust and The Moxie Foundation.

Author Contributions J.W., D.O. and J.C.I.B. conceived the study. J.W. and D.O. derived mESC, EpiSC and rsEpiSC lines. J.W., D.O. and C.R.E. designed and performed *in vivo* embryo grafting experiments. J.W., D.O., M.L., K.S., L.M., Z.L., T.H. and P.R. designed and performed all *in vitro* experiments; J.M.C., J.L. and P.G. helped project design and discussions and performed microarray experiments. I.T., Y.T. performed bisulfite sequencing experiments; M.K. performed teratoma studies; C.L., Y.H., Z.Z., J.R.N. and J.E. performed whole-genome bisulfite sequencing experiments and analysed data. T.D. and B.R. performed ChIP-seq experiments. C.B. and M.L. performed bioinformatics analysis. A.S. analysed global metabolic profiling data; E.A. and N.K. provided technical support. J.W., D.O., M.L. and J.C.I.B. prepared the manuscript.

Author Information Microarray, RNA-seq, ChIP-seq and MethylC-seq data have been deposited in the Gene Expression Omnibus under accession number GSE60605. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C.I.B. (belmonte@salk.edu)

METHODS

No statistical methods were used to predetermine sample size.

Mice. B6.CBA-Tg(*Pou5f1-EGFP*)2Mnn/J³⁷ (The Jackson Laboratory, stock number: 004654) and B6.Cg-Tg(*Prdm1-EYFP*)1Mnz/J³⁸ (The Jackson Laboratory, stock number: 008828) transgenic mice were maintained on the C57BL/6J background. To obtain embryos, ICR or C57BL/6J females were mated with males from ICR, BDF1 (a cross between C57BL/6J and DBA2), B6.CBA-Tg(*Pou5f1-EGFP*)2Mnn/J or B6.Cg-Tg(*Prdm1-EYFP*)1Mnz/J homozygous strains. B6 embryos were obtained by mating female C57BL/6J with male C57BL/6J. Both male and female mice were used at ages between 6 to 25 weeks. No randomization and no blinding were used. All the animal experiments were performed under the ethical guidelines of the Salk Institute, and animal protocols were reviewed and approved by the Salk Institute Institutional Animal Care and Use Committee (IACUC).

Epiblast isolation from post-implantation embryos. Timed-pregnant mice were euthanized for embryo collection at appropriate developmental stages between E5.25 and E7.75. To isolate epiblasts, embryos were first dissected out from decidua. Reichert's membrane, extra-embryonic ectoderm and visceral endoderm were carefully removed mechanically with fine forceps and a tungsten needle. The embryo isolation procedures were performed in media containing DMEM (Gibco), 10% FBS (Hyclone) and 1× penicillin-streptomycin (Gibco). For isolation of epiblast from stage E7.25 and E7.5 embryos, an additional step was taken to mechanically remove the mesodermal layer. Staging of embryos was performed as previously described^{26,39}.

Derivation and culture of mouse rsEpiSC lines. For rsEpiSCs derivation, epiblasts from E5.25 to E7.5 post-implantation mouse embryos were isolated by mechanically removal of Reichert's membrane, visceral endoderm as well as extra-embryonic ectoderm using fine forceps and a tungsten needle. In the case of stage E7.25 and E7.5 embryos, the mesodermal layer was also mechanically removed with a tungsten needle. Isolated epiblasts were placed on MEFs in chemically defined N2B27 medium⁴⁰ supplemented with FGF2 (20 ng ml⁻¹, Peprotech) and IWR1 (2.5 μM, Sigma Aldrich). After 4 day's culture, epiblast outgrowths were dissociated with TrypLE (Life Technologies) and replated onto newly prepared MEFs in one well of a 12-well plate. rsEpiSCs were passaged every 3–4 days with TrypLE at a split ratio of 1:50. For clonal derivation of rsEpiSCs from E6.5 and E7.5 embryos, isolated epiblasts were treated with trypsin-EDTA (0.25%, Life Technologies) for 10 min at 37 °C followed by repeated pipetting (~40 times) with a P200 pipette. Dissociated cells were passed through 40 μm cell strainer to obtain a single-cell suspension and cultured on MEFs in N2B27^{F/R1} media.

Derivation of rsEpiSCs from pre-implantation blastocysts. Timed-pregnant mice were euthanized for blastocyst collection at E3.5. Zona pellucidae were first removed from E3.5 blastocysts after brief treatment with acidic Tyrode's solution (Millipore MR-004-D). ICMs were isolated using immunosurgery. In brief, to remove the trophectoderm layer, blastocysts were incubated with rabbit anti-mouse serum (Sigma-Aldrich) followed by guinea pig complement (Sigma-Aldrich) treatment. The trophectoderm layer was removed by repeated pipetting and ICMs were plated onto MEFs and cultured in N2B27^{F/R1} medium. After 5 day's culture, ICM outgrowths were passaged using TrypLE (Life Technologies) and plated onto newly prepared MEFs.

Derivation and culture of mouse ESC lines. Embryo manipulations were performed under a dissecting microscope (Olympus SZX10). Blastocyst stage embryos were collected from timed-pregnant mice at E3.5 and used for ESC derivation. In brief, zona pellucidae were removed by brief treatment with acidic Tyrode's solution (Millipore MR-004-D). After removing zona pellucidae, embryos were plated on MEF in N2B27^{2iLIF} medium: N2B27 basal medium supplemented with human leukemia inhibitory factor (LIF) (10 ng ml⁻¹, Peprotech), 3 μM CHIR99021 (Selleckchem) and 1 μM PD035901 (Selleckchem). After 6 days in culture, ICM outgrowths were passaged using TrypLE and re-seeded onto newly prepared MEFs for further cultivation. Established mouse ESC lines were cultured either on MEF or poly-L-ornithine (Sigma-Aldrich) and laminin (BD Biosciences) coated plates and passaged every 3–4 days at a split ratio of 1:20.

Derivation and culture of mouse EpiSC lines. E5.75 and E6.5 embryos were used for EpiSCs derivation. In brief, isolated epiblasts were placed onto MEFs plates in EpiSC derivation media: N2B27 basal medium, 20% KnockOut serum replacement (Life Technologies), 20 ng ml⁻¹ Activin-A (Peprotech) and 12 ng ml⁻¹ FGF2 (Peprotech). After 3 days in culture, epiblast outgrowths were passaged as small clumps using collagenase IV (Life Technologies) and replated onto newly prepared MEFs. Established mouse EpiSCs were cultured in EpiSC culture media: N2B27 basal medium, 20% KnockOut serum replacement (KSR), 2 ng ml⁻¹ Activin-A and 12 ng ml⁻¹ FGF2. EpiSCs were cultured on MEF or FBS (Hyclone) coated plates and passaged using collagenase IV every 4–5 days.

Immunofluorescence. For immunofluorescence studies, isolated epiblasts and cells grown on chamber slides (BD Falcon) were fixed with freshly prepared 4% paraformaldehyde in PBS for 15 min at room temperature, and permeabilized/blocked with 1% Triton-X in PBS-contained 10% FBS for 1 h at room temperature. The cells were incubated with primary antibodies in 1% FBS, 0.1% Triton-X in PBS overnight at 4 °C. The next day, cells were washed and incubated with fluorescent-labelled secondary antibodies (Molecular Probes) at 1:500 dilutions for 1 h at room temperature. Cells were washed and mounted in VECTASHIELD with DAPI (Vector Labs). Whole-mount staining of cultured embryos was performed as previously described²⁶. Nuclei were counter stained with DAPI. Specimens were observed and visualized by a Zeiss LSM 780 confocal microscope. Primary antibodies used in this study include: OCT-3/4 (1:200, Santa Cruz, SC-5279), BRACHYURY (T) (1:300, R&D, AF2085), SSEA-1 (1:50, DSHB, MC480), SSEA-4 (1:50, DSHB, MC-813-70), SOX2 (1:100, EMD Millipore, AB5603), NANOG (1:100, EMD Millipore, SC1000), FOXA2 (1:50, Santa Cruz, SC-6554), TUJ1 (1:1000, Sigma-Aldrich, T2200), α-SMA (1:600, Sigma-Aldrich, A5228), H3K27ME3 (1:100, Abcam, ab6002), TRA-1-60 (1:50, Santa Cruz, SC-21705), TRA-1-80 (1:50, Santa Cruz, SC-21706), DNMT3B (1:50, Santa Cruz, SC-10236).

RNA FISH. Mouse *Xist* probes with Quasar 570 dye were purchased from Biosearch Technologies (SMF-3011-1). FISH hybridization was performed following manufacture's protocol (https://www.biosearchtech.com/assets/bti_stellaris_protocol_adherent_cell.pdf). Specimens were observed and visualized by a Zeiss LSM 780 confocal microscope.

Single-cell cloning assay. For mouse cell lines, 500 cells were seeded into 12-well plates on MEFs and cultured in mESC culture medium (N2B27^{2iLIF}), EpiSC culture medium (N2B27^{KSR+F/A} with and without 10 μM Y-27632 treatment) and rsEpiSCs culture medium (N2B27^{F/R1}), respectively. Five days after seeding, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, and colonies were visualized by alkaline phosphatase staining (Vector Laboratories) and/or OCT4 (Santa Cruz, sc-5279) immunohistochemistry (DAB, sigma, D3939). For human cell lines, 500 cells were seeded into Matrigel-coated 12-well plates and cultured in mTeSR1 medium or human rsESC culture medium. Six days after seeding, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, and colonies were visualized by alkaline phosphatase staining. For rhesus macaque cell lines, 500 cells were seeded into 12-well plates on MEFs and cultured in CDF12 medium or rhesus macaque rsESC culture medium. Six days after seeding, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, and colonies were visualized by alkaline phosphatase staining and OCT4 immunohistochemistry.

Epiblast grafting and *in vitro* embryo culture. E7.5 mouse embryos (ICR) were dissected out from decidua. Reichert's membrane, the parietal endoderm as well as the majority of the trophoblast layer, which is part of the parietal yolk sac, were completely removed from the embryos with fine forceps, resulting in a non-intact and non-viable embryo. Grafting cells into the non-intact embryo epiblast was performed manually with an aspirator tube assembly (Drummond) and a hand-pulled glass capillary (Drummond, Microcaps, 50 μl). Before grafting, cells were washed twice with PBS. The cells for grafting were scratched off culture plates using a 20 μl pipette tip, and then cut into small pieces containing 40–50 cells using a tungsten needle. The embryo was held loosely by forceps, and the pulled glass capillary was inserted into the indicated regions of the epiblast. A small volume of dissection medium was expelled out from the tip of the capillary to make an opening in the epiblast and sections of the epiblast/ectoderm, mesoderm and/or endoderm cells were removed from the embryo to further ensure the embryo is in a non-intact and non-viable status prior to grafting. A clump of cells was gently placed inside the opening and the glass capillary was slowly drawn out of the embryo. Injected non-intact embryos were applied to *in vitro* embryo culture in 50% commercial rat serum (Harlan, B.4520) as described previously⁴¹. After 36 h, cultured embryos were washed twice with PBS and fixed in 4% PFA overnight at 4 °C and subsequently used for immunohistochemical analysis, as described above. Kusabira-Orange-labelled mouse EpiSCs and rsEpiSCs, GFP-labelled human H9 ESC, H9 rESC, and GFP-labelled rhesus macaque rsESC line ORMES23 were used for the grafting experiments. Please refer to Supplementary Fig. 1 for an illustrated diagram for the epiblast grafting procedure. The Wisconsin stem cell lines are not permitted for research involving mixing of Wisconsin Materials with an intact embryo, either human or non-human; implanting Wisconsin Materials or products of the Wisconsin Materials in a uterus; or attempting to make whole embryos with Wisconsin Materials by any method. Therefore, grafting of Wisconsin stem cell lines were performed only on non-intact, non-viable post-implantation mouse embryos *in vitro*. The experiments were approved by the Salk Institute embryonic stem cell research oversight committee.

Culture of primate PSCs and rsPSCs. Human ESC lines H1 (WA01) and H9 (WA09) were obtained from WiCell and authenticated by short tandem repeat (STR) profiling. Human and chimpanzee PSCs were cultured either on MEFs in

CDF12 media containing DMEM/F12 (Life Technologies, 11330-032), 20% KnockOut serum replacement (Life Technologies, 10828), 2 mM Glutamax (Life Technologies, 35050-061), 0.1 mM NEAA (Life Technologies, 11140-050), 0.1 mM β -mercaptoethanol (Gibco, 21985) and 4 ng ml⁻¹ FGF2 (Peprotech), or on plates pre-coated with Matrigel (BD Biosciences) using mTeSR1 media⁴². Rhesus macaque PSCs were cultured on MEFs in CDF12 media. Conventional primate PSCs were passaged every 4–5 days either using collagenase IV (Life Technologies) (MEF) or Dispase (Sigma) (Matrigel) at a split ratio of 1:5. Human and chimpanzee rPSCs were cultured either on MEFs or on plates pre-coated with Matrigel (BD Biosciences) using a customized mTeSR1 medium, where an mTeSR1 base medium⁴² lacking FGF2 and TGF β 1 was made in-house and was supplemented with 20 ng ml⁻¹ FGF2 and 2.5 μ M IWR1 to complete. rPSCs were passaged every 4–5 days with TrypLE (Life Technologies) at a split ratio of 1:10. Rhesus macaque rPSCs were cultured on MEF or plates pre-coated with FBS or Matrigel in a modified N2B27 medium: DMEM/F12 (Life Technologies, 11330-032) and Neurobasal medium (Life Technologies, 21103-049) mixed at 1:1 ratio, 1 \times N2 supplement (Life Technologies, 17502-048), 1 \times B27 supplement (Life Technologies, 17504-044), 2 mM Glutamax (Life Technologies, 35050-061), 0.1 mM NEAA (Life Technologies, 11140-050), 0.1 mM β -mercaptoethanol (Gibco, 21985) and 2 mg ml⁻¹ BSA (Sigma), supplemented with FGF2 (Peprotech, 20 ng ml⁻¹) and IWR1 (Sigma, 2.5 μ M). Rhesus macaque rPSCs were passaged every 4–5 days with TrypLE at a split ratio of 1:10. Tests for mycoplasma contamination were routinely performed for all the cell lines using PCR-based approach or MycoAlert mycoplasma detection kit (Lonza) following the manufacturer's recommendation every ten passages.

Flow cytometry analysis. For intracellular FACS analysis, cells were first stained for cell surface markers and then fixed and permeabilized using the BD Cytofix/Cytoperm kit for staining of intracellular antigens. For cell-cycle analysis, human cells were fixed and permeabilized using the BD Cytofix/Cytoperm kit, stained with Alexa Fluor 647 anti-human Ki-67 (Biolegend) and DAPI and then analysed on a BD LSRFortessa cytometer. Mouse cells were fixed in cold 70% ethanol, stained with propidium iodide (ebioscience) and then analysed on a BD LSRFortessa cytometer. Other antibodies used for FACS analysis were: Stemgent StainAlive DyLight 488 anti-Human TRA-1-60 Antibody (09-0068), Stemgent StainAlive DyLight 488 Mouse IgM, κ Isotype Control (09-0072), R&D Systems anti-human/Mouse Oct-3/4 Allophycocyanin MAb (IC1759A), R&D Systems human/Mouse SSEA-4 Phycoerythrin MAb (FAB1435P), eBioscience anti-Human/Mouse SSEA-1 eFluor 660 (50-8813-41), Biogend Alexa Fluor 488 anti-Tubulin Beta 3 (TUBB3) antibody (118213), Abcam anti-NANOG antibody (ab80892) and BD Pharmingen mouse anti-mouse NANOG antibody (560259). **Metabolic flux analysis.** Seahorse bioscience extracellular flux (XF96) analyser was used to measure oxygen consumption rate and extracellular acidification rate of EpiSCs and rsEpiSCs. Cells were plated in XF96 Cell Culture Microplates (Seahorse bioscience, no. 101084-004) pre-coated with FBS at a density of 3–4 \times 10⁴ per well. The next day cells were treated with XF Cell Mito Stress Test Kit (10 μ g ml⁻¹ Oligomycin, 1 μ M FCCP, 1 μ M antimycin + 1 μ M rotenone) or XF Glycolysis Stress Test Kit (10 mM Glucose, 10 μ g ml⁻¹ Oligomycin, 10 mM 2 Deoxy-D-glucose) and measured following the manufacturer's instructions.

Primordial germ cell induction. PGC induction from mESCs was performed following a published protocol⁴³. In brief, mESCs maintained in N2B27^{21,41F} media on poly-L-ornithine (Sigma-Aldrich) and laminin (BD Biosciences) coated plates were trypsinized, counted and about 1 \times 10⁵ cells were seeded into one well of a 12-well plate pre-coated with human plasma fibronectin (15.7 μ g ml⁻¹, Millipore) in N2B27 medium containing 20 ng ml⁻¹ Activin-A (Peprotech), 12 ng ml⁻¹ FGF2 (Peprotech) and 1% KnockOut serum replacement (Life Technologies). The medium was changed daily. After 2 days in culture, PGC-Like Cells (PGCLCs) were induced by plating 1.0 \times 10³ cells per well of low-cell-binding U-bottom 96-well plate (NUNC) in GK15 medium containing GMEM (Sigma), 15% KnockOut serum replacement, 0.1 mM sodium pyruvate, 0.1 mM NEAA, 0.1 mM β -mercaptoethanol, 2 mM Glutamax and 100 U ml⁻¹ penicillin and 0.1 mg ml⁻¹ streptomycin supplemented with BMP4 (500 ng ml⁻¹; R&D Systems), BMP8B (500 ng ml⁻¹; R&D Systems), LIF (10 ng ml⁻¹; Peprotech), SCF (100 ng ml⁻¹; Peprotech), and EGF (50 ng ml⁻¹; Peprotech).

DNA constructs. For packaging of lentiviral vectors, pMDLg/prRE, pRSV-Rev and pMD2.G plasmids were purchased from Addgene (12251, 12253 and 12259). To assess the efficiency of targeted mutagenesis at the *LRK2* locus, we purchased CAS9 expression (hCas9) and guide RNA cloning (gRNA_Cloning Vector) plasmids from Addgene (41815 and 41824)⁴⁴. To construct the mCherry expression gRNA cloning vector (pCAGmCherry-gRNA), the CAG promoter driven mCherry expression cassette was subcloned into the gRNA cloning vector. The *LRK2* target chosen (AGATTCTTTAGATAGTCTAG) is 20 bp in length, has a NGG protospacer adjacent motif (PAM) sequence at the downstream position, and was subcloned into pCAGmCherry-gRNA as per the following protocol (<http://www.addgene.org/static/data/93/40/adf4a4fe-5e77-11e2-9c30-003048dd6500.pdf>)

using the following primers (5'-TTTCTGGCTTTATATATCTTGTGGAAG GACGAAACACCGGATTCTTTAGATAGTCTAG-3' and 5'-GACTAGCCTTA TTTTAACCTTGCTATTCTAGCTCTAAACCTAGAGTATCTAAAGAAT CC-3'). To assess the frequency of homologous recombination mediated gene targeting, we purchased pCas9_GFP (Addgene 44719), pEGIP*35 (Addgene 26776) and tGFP (Addgene 26864). To construct the mutated GFP target gRNA expression vector, the mutated GFP target sequence (CAGGTAATCTCGA GAGCTT) was subcloned into the gRNA_Cloning Vector as described above using the following primers (5'-TTTCTGGCTTTATATATCTTGTGGAAGGACG AAACACCGAGGGTAATCTCGAGAGCTT-3' and 5'-GACTAGCCTTATTTT AACTTGCTATTCTAGCTCTAAACCAAGCTCTCGAGATTACCTC-3'). TALENs recognizing the target site were constructed using the Golden Gate Assembly method with the TALE Toolbox kit from Addgene (cat. no.1000000019)⁴⁵. The constructed TALEN pair targeting the mutated GFP gene was named TALEN-L and TALEN-R.

Generating of a mutant eGFP human ESC reporter. To assess the efficiency of gene targeting in human ESCs and rsESCs, we established a mutated GFP gene-based reporter system, similar to one previously described⁴⁶. In brief, pEGIP*35 was cotransfected with pMDLg/prRE, pRSV-Rev and pMD2.G, and packaged and purified as lentiviral vectors according to a published protocol⁴⁷. H9 human ESCs cultured on Matrigel were incubated with 10 μ M Y-27632 overnight and then individualized with Accumax (Innovative Cell Technologies). Cells were transduced in suspension with lentiviral EGIP*35 vector in the presence of Y-27632 and 4 μ g ml⁻¹ polybrene for 1 h. After brief centrifugation to remove any residual lentiviral vector, the cells were seeded on irradiated DR4 MEF feeders (ATCC) in CDF12 media containing 10 μ M Y-27632. Three days after transduction, puromycin (1 μ g ml⁻¹; Invitrogen) was added to the medium. After 2 weeks, ESC colonies were manually picked onto fresh MEF feeders and expanded as mutant eGFP reporter human ESC lines. To generate corresponding rsESCs, the mutant eGFP reporter ESCs were first converted to rsESCs after culturing in human rsESC medium for five passages before lentiviral infection.

Measurement of targeted mutagenesis with CRISPR/Cas9 in human ESCs and rsESCs. To compare the targeted mutagenesis efficiency in human H1 ESCs and rsESCs, 1.5 \times 10⁷ feeder-free cultured cells were dissociated by TrypLE (Invitrogen), and resuspended in 1 ml of medium with or without 10 μ M ROCK inhibitor Y-27632 for ESCs or rsESCs, respectively. These cells were electroporated with 20 μ g of Cas9-2A-GFP expression vector (pCas9_GFP) and 20 μ g of *LRK2* target mCherry-gRNA expression vectors, and were plated onto 100-mm dishes pre-coated with Matrigel. Two days after electroporation, the cells were dissociated by TrypLE, and Cas9 and gRNA expression cells were sorted out as eGFP/mCherry double-positive cells by BD influx cell sorter (BD), and ~10,000 cells were plated onto 100-mm dishes pre-coated with MMC-treated MEFs. Two weeks later, visible colonies were counted and each ~96 colonies were transferred to a 96-well plate and genomic DNA was extracted following a previous report⁴⁶. To determine targeted mutant clones, the target *LRK2* site was PCR-amplified with the following primers: forward 5'-AGTCTCCAAAATTGGGCTTTGCCTGAGATAG ATTTGTC-3' and reverse 5'-CCAGTTTCTATTGGTCTCCTTAACCTGT-3' with PrimeSTAR GXL DNA Polymerase (TAKARA) following the manufacturer's protocol. Amplicons were sequenced using an ABI 3730 sequencer (Applied Biosystems) with reverse primer.

Measurement of gene targeting efficiencies with CRISPR/Cas9 and TALEN in human ESCs and rsESCs. To compare the targeted mutagenesis efficiency in human mutant eGFP reporter ESCs and rsESCs, 2 \times 10⁵ feeder-free cultured cells were dissociated by TrypLE (Invitrogen), and plated into 1 well of a 6-well plate with (ESCs) or without (rsESCs) 10 μ M ROCK inhibitor Y-27632. The following day, the cells were transfected with a total of 2 μ g of DNA using FuGENE HD (Promega). For CRISPR/Cas9-mediated gene targeting, 0.5 μ g of Cas9 expression vector (hCas9), 0.5 μ g of mutated GFP target gRNA expression vector and 1 μ g of donor vector (tGFP) were co-transfected. For TALEN-mediated gene targeting, 0.5 μ g of TALEN-L, 0.5 μ g of TALEN-R and 1 μ g of donor vector (tGFP) were co-transfected. Five days after transfection, GFP-positive cells were detected by BD LSRFortessa, and the gene-targeting frequencies per 5 \times 10⁵ cells were determined.

Human iPSC generation. Reprogramming of human fibroblasts with episomal vectors was performed as previously described⁴⁷ with minor modifications. Episomal plasmids pCXLE-EGFP (27082), pCXLE-hOCT3.4-shp53-F (27077), pCXLE-hSK (27078) and pCXLE-hUL (27080) were obtained from Addgene. 2 \times 10⁶ human foreskin fibroblasts (HFF, ATCC, CRL-2429) or BJ fibroblasts (ATCC CRL-2522) were nucleofected with the episomal vectors using 4D-Nucleofector (Lonza) using P2 Primary Cell 4D-Nucleofector kit (Lonza, V4XP). Five days post-nucleofection fibroblasts were replated onto mitotically inactivated MEFs. The next day, medium was changed to human ESC culture media or human

rsESC culture medium. Putative iPSC colonies were picked between day 24 and day 32 and transferred to newly prepared MEFs. For evaluating the efficiency and colony quality of iPSCs, on day 25 cells were fixed with 4% PFA for 15 min at room temperature and stained for alkaline phosphatase activity with Vectastain ABC-AP kit (Vector Laboratories).

RNA preparation and real-time PCR. Total RNAs were extracted by using the QIAGEN RNeasy mini kit or the micro kit according to the manufacturer's instructions. RNAs were reverse-transcribed using iScript RT Supermix (Bio-Rad), and real-time PCR was performed using SsoAdvanced Universal SYBR Green Supermix in CFX384 (Bio-Rad). Expression levels of each gene were normalized to GAPDH (mouse) and HPRT (human) expression and calculated using comparative C_T method.

Bisulfite sequencing. Genomic DNA was purified using DNeasy kit (Qiagen). Bisulfite conversion of genomic DNA was carried out using the Zymo EZ DNA Methylation-direct Kit (Zymo Research). *Oct4*, *stella* and *Dppa5a* promoter regions were amplified by EpiTaq HS (Takara) under the nested PCR condition: 1st PCR, 98 °C for 30 s, 35 cycles (98 °C for 10 s, 55 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min; 2nd PCR, modified to 25 cycles. The sequences of PCR primers are described in Supplementary Table 8. PCR products were cloned into the pCR2.1-TOPO vector (Invitrogen) and sequenced. Sequence data was analysed using QUMA (<http://quma.cdb.riken.jp>).

DNA microarray and data analysis. Total RNA of all samples was extracted using TRIzol Reagent (Invitrogen) and purified by RNeasy Mini Kit (QIAGEN). Affymetrix Mouse Gene 2.0 ST Gene Expression Arrays were performed by the Genomics Core Facility at the Center for Regenerative Medicine in Barcelona according to the manufacturer's protocol (Affymetrix). Gene expression clustering was performed using Cluster 3.0 and visualized using Java TreeView. PCA analysis was performed using R (<http://www.r-project.org>) and visualized in 3D using the rgl library.

RNA-seq and data analysis. RNA-seq libraries were sequenced on an Illumina HiSeq 2500 according to the manufacturer's instructions. Reads were aligned to the human genome (hg19, GRCh37) using STAR (PMID: 23104886). RNA-seq alignments were normalized to the total number of aligned reads and visualized by using HOMER (<http://homer.salk.edu/homer/>)⁴⁸ to generate custom tracks for the UCSC Genome Browser (<http://genome.ucsc.edu/>). Gene expression values were generating for RefSeq annotated transcripts using HOMER and differential expression calculations were performed using EdgeR⁴⁹. Gene Ontology analysis was performed using DAVID (<http://david.abcc.ncifcrf.gov/>). GSEA analysis was performed using the GSEA software with default parameters and permutation number set to 100. Gene expression clustering was performed using Cluster 3.0 and visualized using Java TreeView. PCA analysis was performed using R and visualized in 3D using the rgl library. Spearman's rank correlation matrix was used to compare rsEpiSCs with *in vivo* isolated four quadrants of the E6.5 epiblasts (anterior-proximal, anterior-distal, posterior-proximal and posterior-distal). Genes were selected from the RNA-seq data sets if the standard deviation of their fragments per kilobase of exon per million reads mapped (FPKM) values among the four *in vivo* epiblast samples is greater than 25% of the mean.

Chromatin Immunoprecipitation. ChIP experiments were carried out as described previously⁵⁰ with modifications. In brief, cells were fixed with 1% formaldehyde at 37 °C for 10 min and then quenched with glycine at 37 °C for 5 min. Fixed cells were sonicated using Epishear (Active Motif) to achieve 200–700 bp size chromatin fragments. Solubilized chromatin was immunoprecipitated with antibody against H3K4me3 (Abcam 8580) and H3K27me3 (Millipore 07-449). Antibody–chromatin complexes were pulled down using Dynabeads protein A (Invitrogen), washed and then eluted. After cross-linking reversal, RNase and proteinase K treatment, immunoprecipitated DNA was purified using AMPure beads (Beckman Coulter).

Library preparation and Illumina sequencing. ChIP DNA were end-repaired and 5' phosphorylated using T4 DNA Polymerase, Klenow and T4 Polynucleotide Kinase (Enzymatics). A single adenine was added to 3' ends by Klenow (3→5' exo-), and double-stranded Bioo Illumina Adapters (Bioo Scientific) were ligated to the ends of the ChIP fragments. Adaptor-ligated ChIP DNA fragments were subjected to 15 cycles of PCR amplification using Q5 polymerase (NEB). AMPure beads were used to purify DNA after each step (Beckman Coulter). Pooled libraries were sequenced on the NextSeq500 for single-end 75 bp using high-output flowcell according to the manufacturer's instructions. Reads were aligned to the reference genome (hg19, GRCh37) by using the program bowtie2 with default parameters. Mapped reads were then investigated for the presence of enrichment against the input. Peaks with a false-discovery rate lower or equal to 0.05 were kept for the further analysis. BEDtools package was used for detecting the Ensembl genes (version 70) overlapping with the detected peaks.

Preparation of methylC-seq libraries and sequencing. MethylC-seq libraries were prepared as previously described^{22,51,52}. The only significant modification of the procedure was that the library amplification was performed with KAPA HiFi HotStart Uracil + ReadyMix (Kapa Biosystems KK2802) using the following PCR conditions: 2 min at 95 °C, 30 s at 98 °C, 4 cycles of (15 s at 98 °C, 30 s at 60 °C, 1 min at 72 °C), and 10 min at 72 °C. Libraries were sequenced on an Illumina HiSeq 2500 up to 101 cycles.

Processing of methylC-seq data and DMR calling. MethylC-seq reads were processed with MethylPy pipeline. (<https://bitbucket.org/schultzmatt/methylpy/>)⁵³. Bowtie index for methylome mapping was constructed using the *build_ref* function imported from the *methylpy.call_mc* library. The mapping of methylC-seq reads were performed using the *run_methylation_pipeline* function imported from *methylpy.call_mc* library. The identification of differentially methylated regions (DMRs) was performed using the *methylpy.DMRfind*. *DMRfind* function imported from the *methylpy.DMRfind* library with a FDR cutoff of 0.01 for calling differentially methylated sites (DMSs). DMSs with methylation changes in the same direction were combined into DMRs if they were located within 250 bp of one another. DMRs containing less than four DMSs were discarded. DMRs showing consistent hyper- or hypomethylation states in biological replicates as determined by MethylPy were used for further analyses.

Bioinformatics of analysis of methylC-seq data. UMRs and LMRs were identified using MethylSeekR⁵⁴ with $m = 0.5$ and a FDR cutoff of 0.05. The list of CpG islands was downloaded from the UCSC genome browser for the mm10 reference genome. Transcription start sites were defined by GENCODE M2 transcripts annotation. Gene Ontology analysis of rsEpiSCs hyper-DMRs was performed with GREAT with the default setting of 'basal plus extension' method⁵⁵.

Extraction and metabolomics. Prepare 80:20 methanol:water solution (estimating 1 ml per sample), cool to –80 °C (4–16 h). Prepare another batch of 80:20 methanol:water solution (estimating 1 ml per sample), cool to 4 °C (overnight is preferred, but a few hours should be fine also). Thaw cell pellets on ice (approximately 20 min) and remove all residual PBS from top of cell pellets. Add 1 ml of –80 °C 80:20 methanol:water solution (keep this solution on dry ice to keep it cold) to cell pellet, mix and leave on dry ice for 15 min. Centrifuge at 2,000g, 5 min, 4 °C. Transfer supernatant onto precooled 4 ml glass vial on ice, set aside on ice. Add 0.5 ml 4 °C 80:20 methanol:water solution to cell pellet, mix and leave on ice for 15 min. Centrifuge at 2,000g, 5 min, 4 °C. Transfer supernatant into previous glass vial containing extract on ice, set aside on ice. Repeat with another 0.5 ml 4 °C 80:20 methanol:water, combine all supernatants (approximately 2 ml total volume). Dry under a gentle stream of nitrogen, flush each sample briefly with nitrogen, cap and store lipids at –80 °C and reconstitute in 1:1 methanol:water for liquid chromatography–mass spectrometry.

Samples were analysed using a 15 cm SeQuant EMD Millipore ZIC pHILIC column (15 cm, 5 µm particle size, 2.1 mm inner diameter (ID)) at a flow rate of 0.100 ml per minute. Mobile phases A and B were 20 mM ammonium carbonate with 0.1% v/v ammonium hydroxide and acetonitrile, respectively. The mobile phase composition started at 100% B and decreased to 40% B over 20 min. Mobile phase B was then raised to 20% over 0.1 min and maintained at 20% B for 4.9 min for washing out the most strongly retained hydrophilic metabolites. The column was then allowed to re-equilibrate to starting conditions over 0.1 min of 100% B and kept there for the next 11.9 min. Samples were run on the Bruker Impact HD q-TOF mass spectrometer with internal calibration of each liquid chromatography–mass spectrometry run using sodium formate clusters at the beginning and end of the run. The instrument was carefully tuned to transmit low m/z ions.

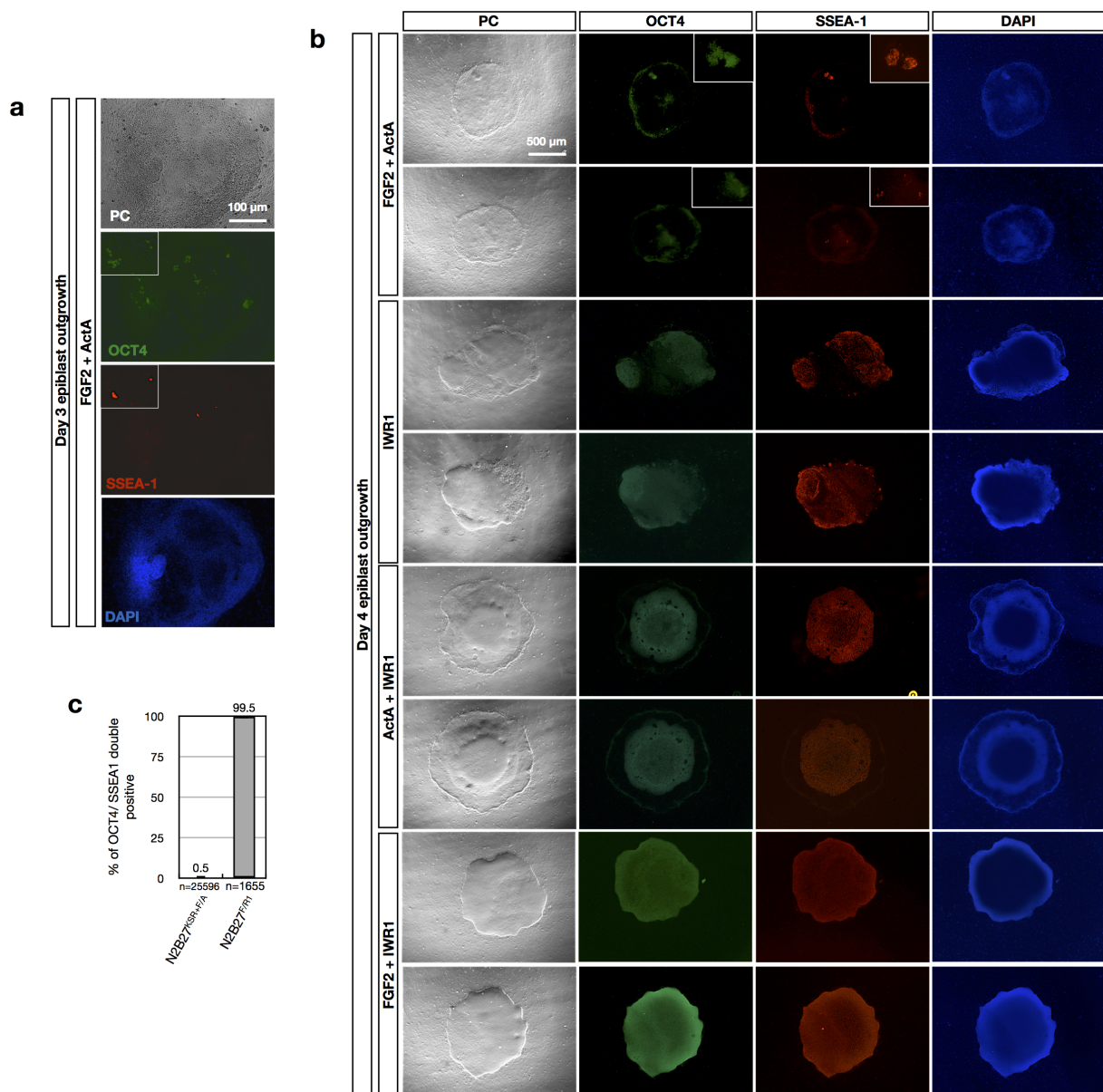
Extraction and lipidomics. Harvest cells by scraping into cold PBS on ice, pellet cells at 1,400g, 3 min, 4 °C. Suspend cell pellet in small amount (50–100 µl) of PBS and transfer rest of cells into Teflon-capped glass vials on ice. Add aqueous buffer to 1 ml. Add 1 ml methanol followed by 2 ml chloroform. Shake vigorously for 30 s, then vortex on high for 15 s (room temperature). Centrifuge at 2,200g, 6 min, 4 °C. Gently remove chloroform (bottom) layer and transfer the chloroform into a clean vial, transfer again from the first clean vial into a second vial to remove additional aqueous contaminants. Dry chloroform under a gentle stream of nitrogen (30–45 min), flush each sample briefly with nitrogen, cap and store lipids at –80 °C and reconstitute in chloroform for liquid chromatography–mass spectrometry.

A Gemini C18 reversed phase column (5 µm, 4.6 × 50 mm, Phenomenex) and a C18 reversed phase guard column (3.5 µm, 2 × 20 mm, Western Analytical) were used for liquid chromatography–mass spectrometry analysis in negative mode. In positive mode, a Luna C5 reversed phase column (5 µm, 4.6 × 50 mm, Phenomenex) was used together with a C4 reversed phase guard column (3.5 µm, 2 × 20 mm, Western Analytical). 30 µl of each sample was injected using an autosampler. Mobile phase A consisted of a 95:5 water:methanol mixture and mobile phase B consisted of 60:35:5 2-propanol:methanol:water. In negative mode, 0.1% ammonium hydroxide was added to the mobile phases and in

positive mode, 0.1% formic acid plus 5 mM ammonium formate were added. An Agilent 1200 series binary pump was set to a flow rate of 0.1 ml per min for the first 5 min followed by 0.4 ml per min for the remainder of the gradient. At 5 min, concomitant with the increase in flow rate, the gradient was increased from 0% B to 20% B. The gradient increased linearly to 100% B at 45 min, followed by an 8-min wash at 0.5 ml per min with 100% B before re-equilibrating the column with 0% B for 7 min. Mass spectrometry analysis was performed using an Agilent 6220 ESI-TOF fitted with an electrospray ionization (ESI) source. The capillary voltage was set to 3,500 kV and the fragmentor voltage to 100 V. The drying gas temperature was set to 350 °C at a flow rate of 10 l per min with a nebulizer pressure set to 45 p.s.i. Untargeted data were collected using a mass-to-charge range of m/z 100–1,500.

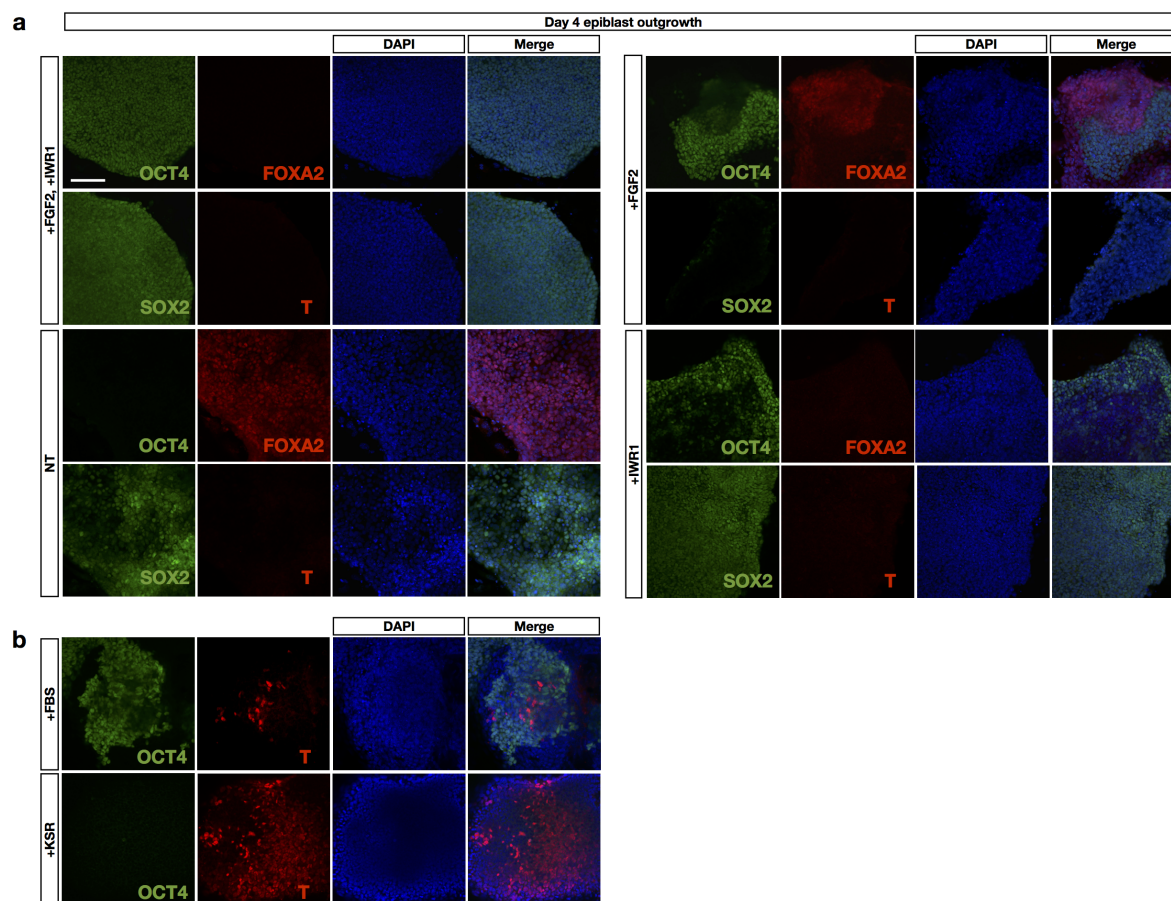
Mass spectrometry data analysis. Data analysis was performed with XCMS online to identify changing metabolites between samples. Samples (EpiSCs and rsEpiSCs) were compared and differences were ranked according to statistical significance as calculated by an unpaired Student's *t* test. For the volcano plot, data were obtained from the 60-min profiling analysis in negative and positive mode and the data were filtered based on retention time range (5–40 min) and abundance ($>5 \times 10^3$ counts). For the metabolite and lipid identification each ion was inspected to ensure that the differences identified by XCMS were reflected in the raw data. Many of the ions in the volcano plots belonged to di-, tri- and tetrapeptides and were not annotated in our final metabolite lists. The metabolite IDs listed are based on metabolites in the METLIN database that are within 5 p.p.m. of the experimental mass measured. This is all done automatically in XCMS online. In several cases, more than one metabolite is identified for a given mass because they all have the same formula. Where required we selected the metabolite that was most appropriate (that is, a mammalian metabolite versus a drug with the same formula) and listed that metabolite.

37. Yoshimizu, T. *et al.* Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41**, 675–684 (1999).
38. Fooksman, D. R., Nussenzweig, M. C. & Dustin, M. L. Myeloid cells limit production of antibody-secreting cells after immunization in the lymph node. *J. Immunol.* **192**, 1004–1012 (2014).
39. Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* **1**, 37–49 (2001).
40. Tong, C., Huang, G., Ashton, C., Li, P. & Ying, Q.-L. Generating gene knockout rats by homologous recombination in embryonic stem cells. *Nature Protocols* **6**, 827–844 (2011).
41. Glanville-Jones, H. C., Woo, N. & Arkell, R. M. Successful whole embryo culture with commercially available reagents. *Int. J. Dev. Biol.* **57**, 61–67 (2013).
42. Ludwig, T. E. *et al.* Feeder-independent culture of human embryonic stem cells. *Nature Methods* **3**, 637–646 (2006).
43. Hayashi, K. & Saitou, M. Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nature Protocols* **8**, 1513–1524 (2013).
44. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
45. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature Biotechnol.* **29**, 149–153 (2011).
46. Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nature Methods* **8**, 409–412 (2011).
47. Li, M. *et al.* Efficient correction of hemoglobinopathy-causing mutations by homologous recombination in integration-free patient iPSCs. *Cell Res.* **21**, 1740–1744 (2011).
48. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
49. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
50. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
51. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
52. Ma, H. *et al.* Abnormalities in human pluripotent cells due to reprogramming mechanisms. *Nature* **511**, 177–183 (2014).
53. Lister, R. *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
54. Burger, L., Gaidatzis, D., Schübeler, D. & Stadler, M. B. Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Res.* **41**, e155 (2013).
55. McLean, C. Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nature Biotechnol.* **28**, 495–501 (2010).
56. Zou, J. *et al.* Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* **5**, 97–110 (2009).
57. Kutner R.H. *et al.* Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. *Nature Protocols* **4**, 495–505 (2009).



Extended Data Figure 1 | The effects of culture parameters on epiblast explants. Isolated epiblasts from E5.75 embryos were plated onto mitotically inactivated MEFs in the following media. **a**, In EpiSC derivation medium containing 20% KnockOut serum replacement (KSR), 20 ng ml⁻¹ Activin-A and 12 ng ml⁻¹ FGF2. Day 3 outgrowth was stained with pluripotency markers OCT4 and SSEA-1. **b**, In N2B27 media supplemented with indicated growth factors and small molecules. After 4 days, outgrowths of plated epiblasts were

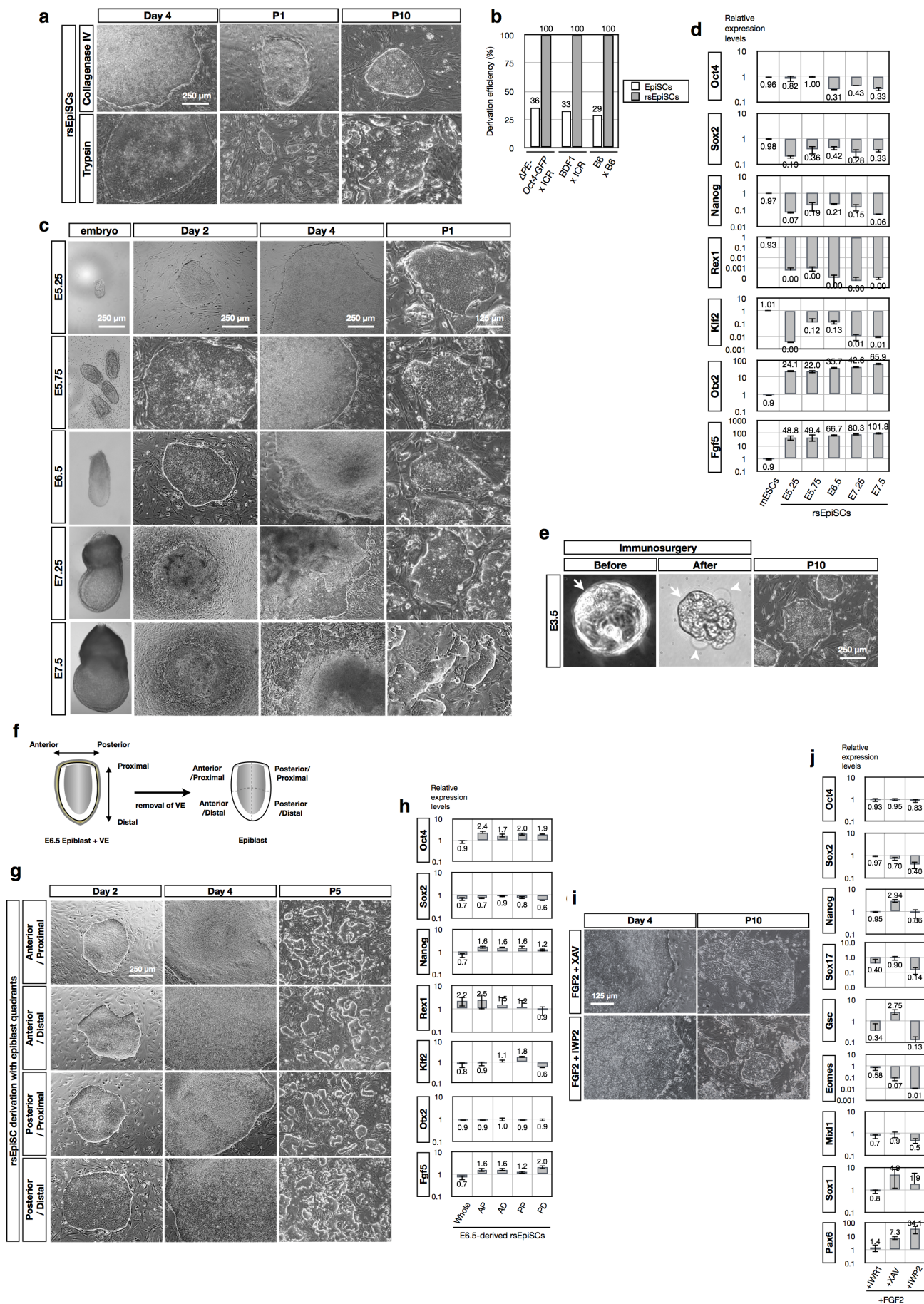
stained for OCT4 and SSEA-1. **c**, Percentages of SSEA-1⁺/OCT4⁺ cells in day 4 epiblast outgrowths in N2B27^{KSR+F/A} and N2B27^{F/R1} culture conditions. A simple randomization method was applied to randomly pick the microscope fields of views for counting the number of SSEA-1⁺/OCT4⁺ cells and total cell numbers. PC, phase contrast. For examining different culture parameter effects, all isolated E5.75 epiblasts were pooled together and randomly allocated to each condition.



Extended Data Figure 2 | Mechanistic studies of rsEpiSC derivation.

Isolated epiblasts from E5.75 embryos were plated onto mitotically inactivated MEFs in the following media. **a**, In N2B27 media with indicated treatments. NT, no treatment. After 4 days, outgrowths of plated epiblasts were stained with endodermal marker FOXA2, mesodermal marker T, neuroectodermal marker

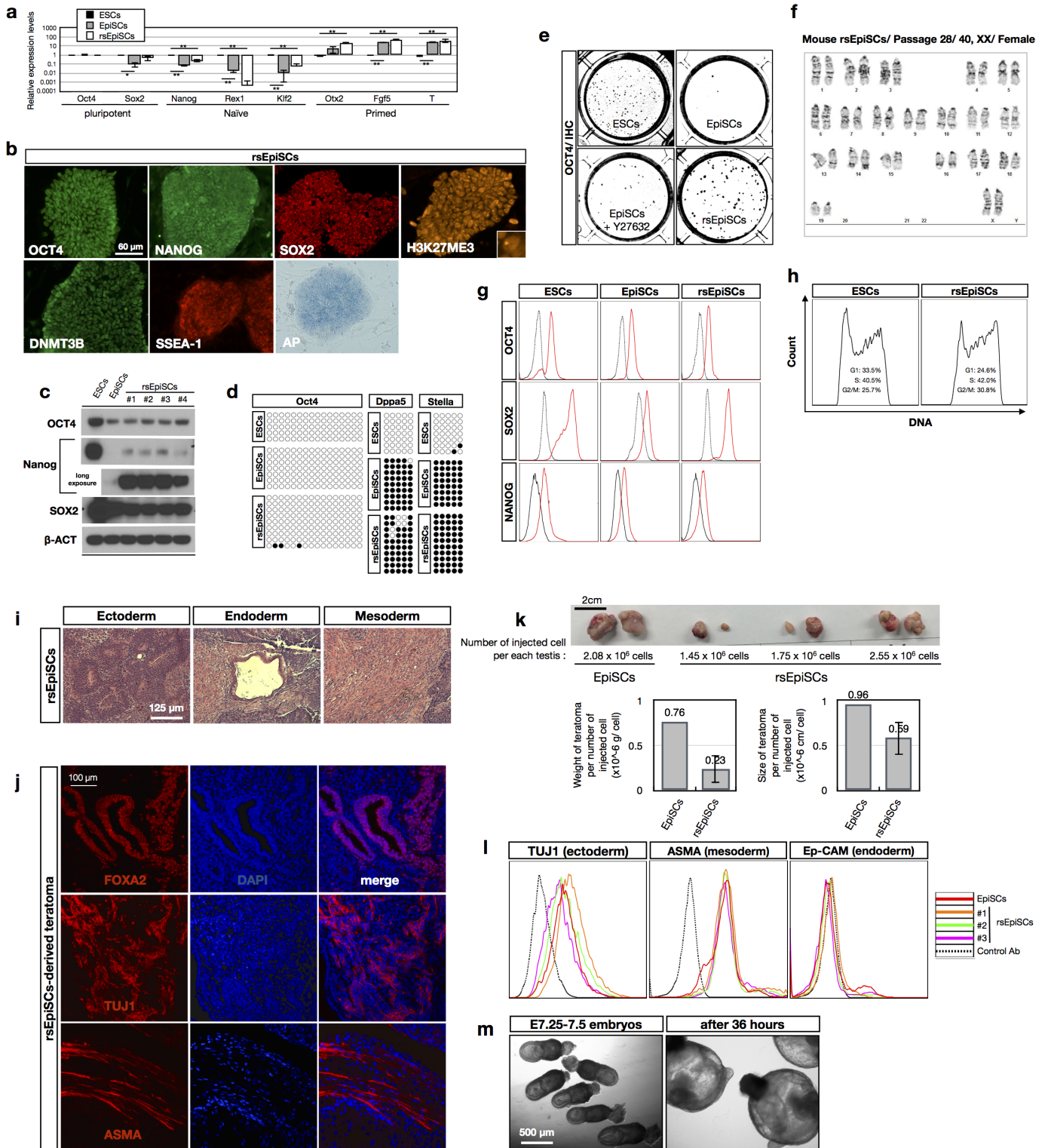
SOX2 and pluripotent marker OCT4. **b**, In N2B27 media supplemented with either 20% FBS (top) or 20% KnockOut serum replacement (KSR; bottom). Day 4 outgrowths were fixed and stained with mesodermal marker T and pluripotency marker OCT4. Nuclei were counterstained with DAPI. Scale bar, 125 μ m.



Extended Data Figure 3 | Highly efficient derivation of rsEpiSCs.

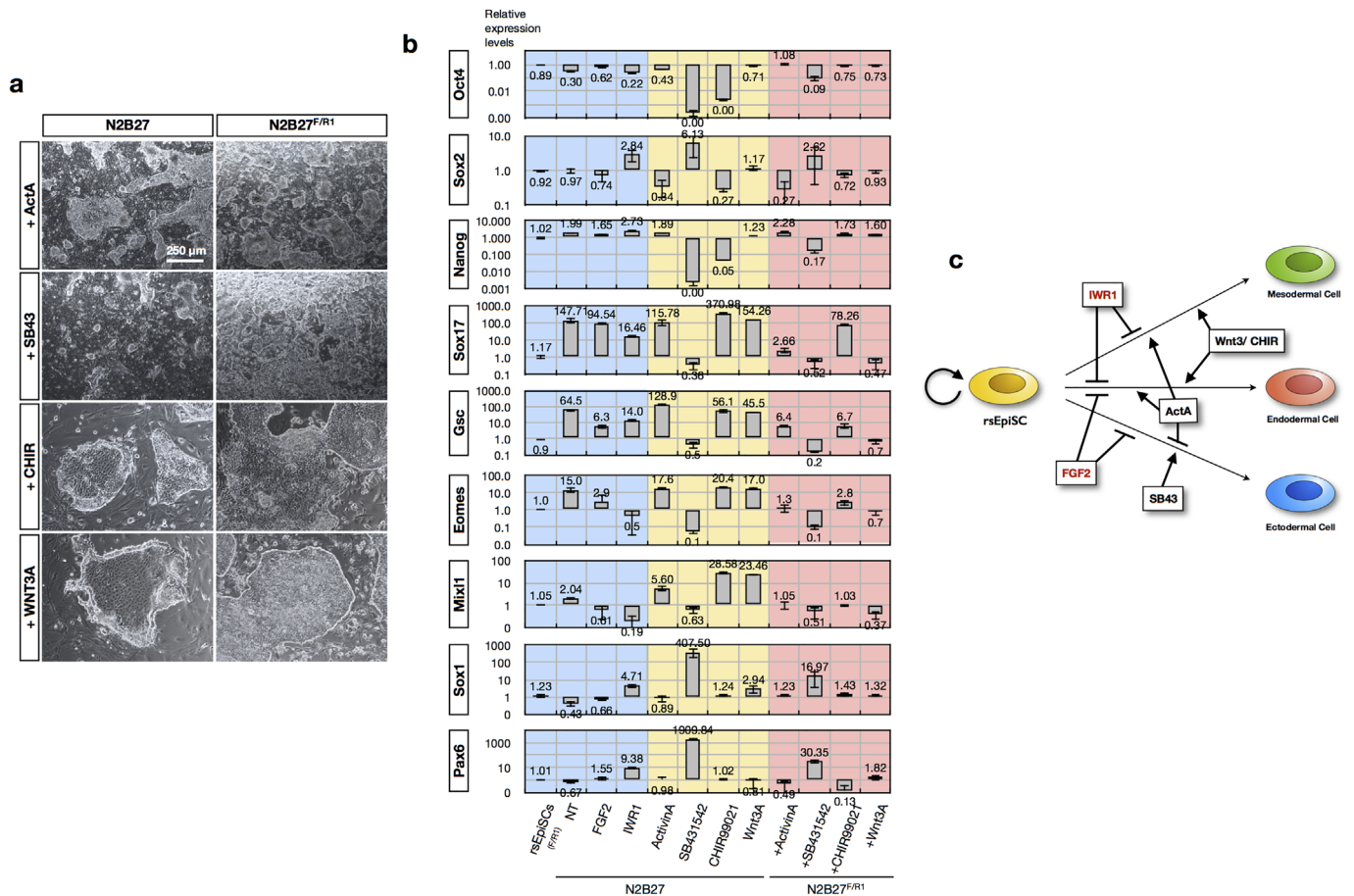
a, Derivation of rsEpiSCs with different passaging methods: collagenase IV (top) and trypsin (bottom). Shown are phase-contrast images of day 4 epiblast outgrowths and cells at passage 1 (P1) and passage 10 (P10). **b**, Derivation efficiency is compared between EpiSCs and rsEpiSCs using isolated E5.75 epiblasts from three different mouse strains. **c**, Derivation of rsEpiSCs using epiblasts isolated from different developmental stages of post-implantation embryos. Typical morphologies of staged embryos at E5.25, 5.75, 6.5, 7.25 and 7.5 are shown in the left panel. Day 2 and day 4 epiblast outgrowths as well as colonies of P1 are shown. **d**, Real-time quantitative PCR analysis of expression of pluripotent (*Oct4*, *Sox2* and *Nanog*), naive (*Rex1* and *Klf2*) and primed (*Otx2* and *Fgf5*) PSC marker genes in mouse ESCs and rsEpiSCs derived from different post-implantation developmental stages. Error bars indicate s.d. ($n = 3$, biological samples). **e**, Derivation of rsEpiSCs from E3.5 pre-implantation blastocysts. Zona pellucida was removed with acidic Tyrode's solution, followed by the removal of trophectoderm by immunosurgery. Isolated inner cell mass was used for the derivation of rsEpiSCs. Arrows and

arrowheads point to the intact trophectoderm and destroyed trophectoderm, respectively. **f**, Schematic representation of dissection of isolated E6.5 epiblasts into four pieces: anterior-proximal (AP), anterior-distal (AD), posterior-proximal (PP) and posterior-distal (PD). **g**, Derivation of rsEpiSCs from four regions of E6.5 epiblasts. Day 2 and day 4 epiblast outgrowths as well as colonies of passage 5 (P5) are shown. **h**, Real-time quantitative PCR analysis of expression of pluripotent (*Oct4*, *Sox2* and *Nanog*), naive (*Rex1* and *Klf2*) and primed (*Otx2* and *Fgf5*) PSC marker genes in rsEpiSCs derived from whole, AP, AD, PP and PD regions of E6.5 epiblasts. Error bars indicate s.d. ($n = 3$, biological samples). **i**, Derivation of rsEpiSCs using other Wnt inhibitors: XAV939 (2.5 μ M) and IWP2 (2.5 μ M). Day 4 epiblast outgrowths and colonies at passage 10 (P10) were shown. **j**, Real-time quantitative PCR analysis of expression of pluripotent (*Oct4*, *Sox2* and *Nanog*), endodermal (*Sox17* and *Gsc*), mesodermal (*Eomes* and *Mixl1*) and neuroectodermal (*Sox1* and *Pax6*) marker genes in rsEpiSCs derived using different Wnt inhibitors. Error bars indicate s.d. (**d**, **h**, **j**, $n = 3$, biological replicates).



Extended Data Figure 4 | Characterization of rsEpiSCs. **a**, Quantitative PCR analysis of expression of pluripotent, naive and primed PSC markers in mouse ESCs, EpiSCs and rsEpiSCs. Error bars indicate s.d. ($n = 3$, biological replicates); t -test, $^{**}P < 0.01$, $^{*}P < 0.05$. **b**, Expression of OCT4, NANOG, SOX2, DNMT3B and SSEA-1 proteins in mouse rsEpiSCs was analysed by immunofluorescence. Mouse rsEpiSCs also displayed weak alkaline phosphatase activity and showed positive H3K27me3 foci confirming an inactivated X chromosome in female rsEpiSCs. **c**, Western blot analyses of OCT4, NANOG and SOX2 protein levels in mouse ESCs, EpiSCs and four different lines of rsEpiSCs. β -actin was used as loading control. For NANOG, an additional long-exposure image was shown (without ESC sample loaded). For full scan associated with **b**, refer to Supplementary Information. **d**, DNA methylation patterns of *Oct4*, *Dppa5* and *Stella* promoters in mESCs, EpiSCs and rsEpiSCs. **e**, Representative bright-field images showing colonies stained by immunohistochemistry for OCT4 expression after being plated at clonal density (500 cells per well), and cultured for 5 days. Y27632 was added at 10 μ M. **f**, Karyotype analysis of mouse rsEpiSCs indicates a normal diploid

chromosome content. **g**, Flow cytometry analysis of OCT4, SOX2 and NANOG expression in mouse ESCs, EpiSCs and rsEpiSCs. **h**, Cell-cycle profiles of mouse ESCs and rsEpiSCs analysed by flow cytometry. **i**, Haematoxylin and eosin staining images of teratomas generated by rsEpiSCs show lineage differentiation towards three germ layers. **j**, Teratomas generated by rsEpiSCs showed tri-lineage differentiation as examined by immunofluorescence analysis using FOXA2 (endoderm), TUJ1 (neuroectoderm) and ASMA (mesoderm) antibodies. **k**, Teratomas generated by injecting indicated number of cells in testis of NOD/SCID male mice. EpiSCs and three different lines of rsEpiSCs were used for comparison. After one month, mice were euthanized. Teratomas were retrieved and measured in size and weight. (EpiSCs, $n = 1$, biological replicate, two technical replicates; rsEpiSCs, $n = 3$, biological replicates, two technical replicates per line; error bars, s.d.) **l**, Flow cytometry analysis of TUJ1, ASMA and Ep-CAM expression in teratomas generated by EpiSCs and three different lines of rsEpiSCs. **m**, Bright-field images of isolated non-intact and non-viable E7.25–7.5 mouse embryos before and after *in vitro* embryo culture.

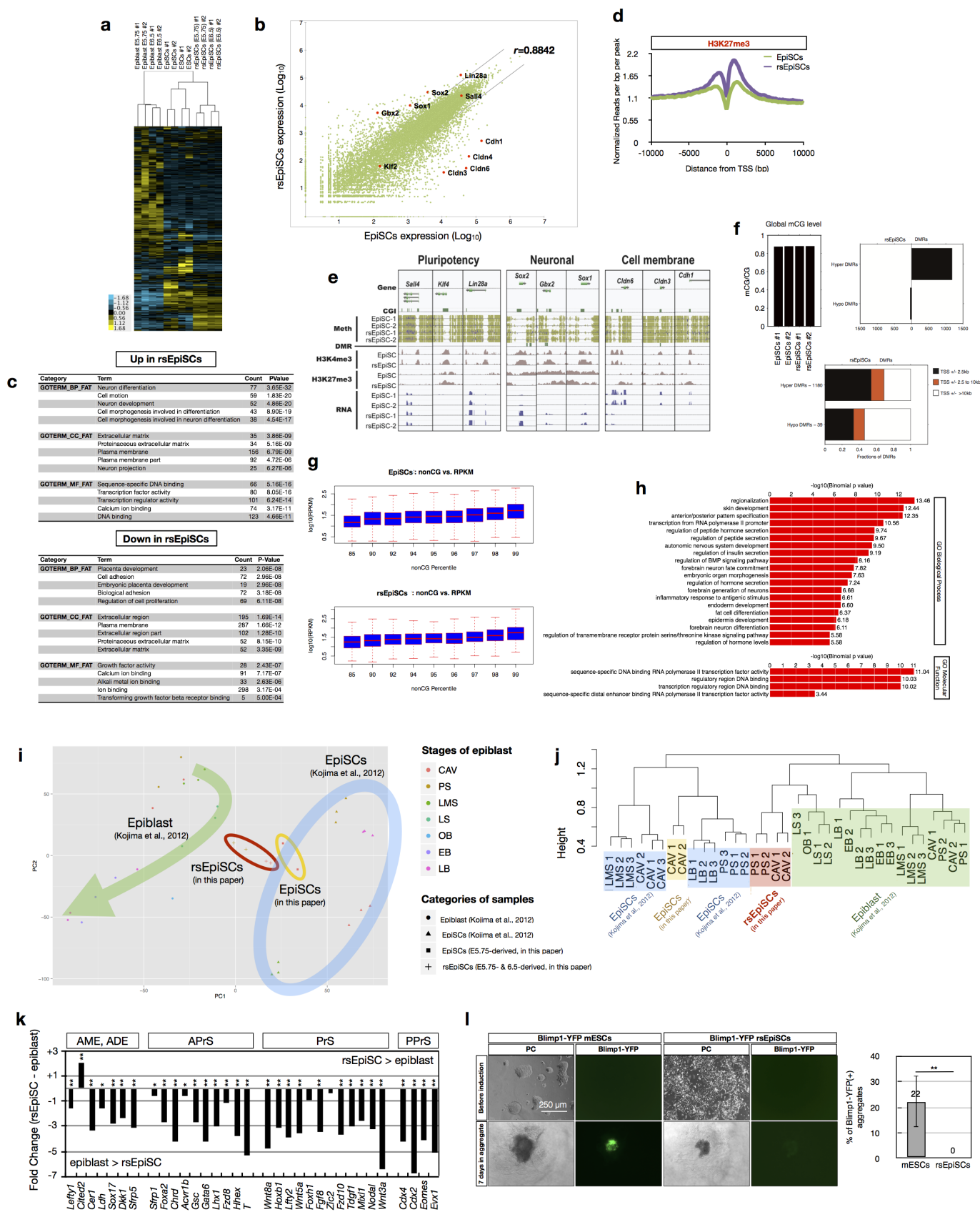


Extended Data Figure 5 | Mechanistic studies of rsEpiSCs self-renewal.

a, Phase-contrast images showing colony morphologies of rsEpiSCs after 4 days of indicated treatments. Left, N2B27 media alone; right, N2B27^{F/R1}.

b, Quantitative PCR analysis of expression of pluripotent (*Oct4*, *Sox2* and *Nanog*), endodermal (*Sox17* and *Gsc*), mesodermal (*Eomes* and *Mixl1*) and

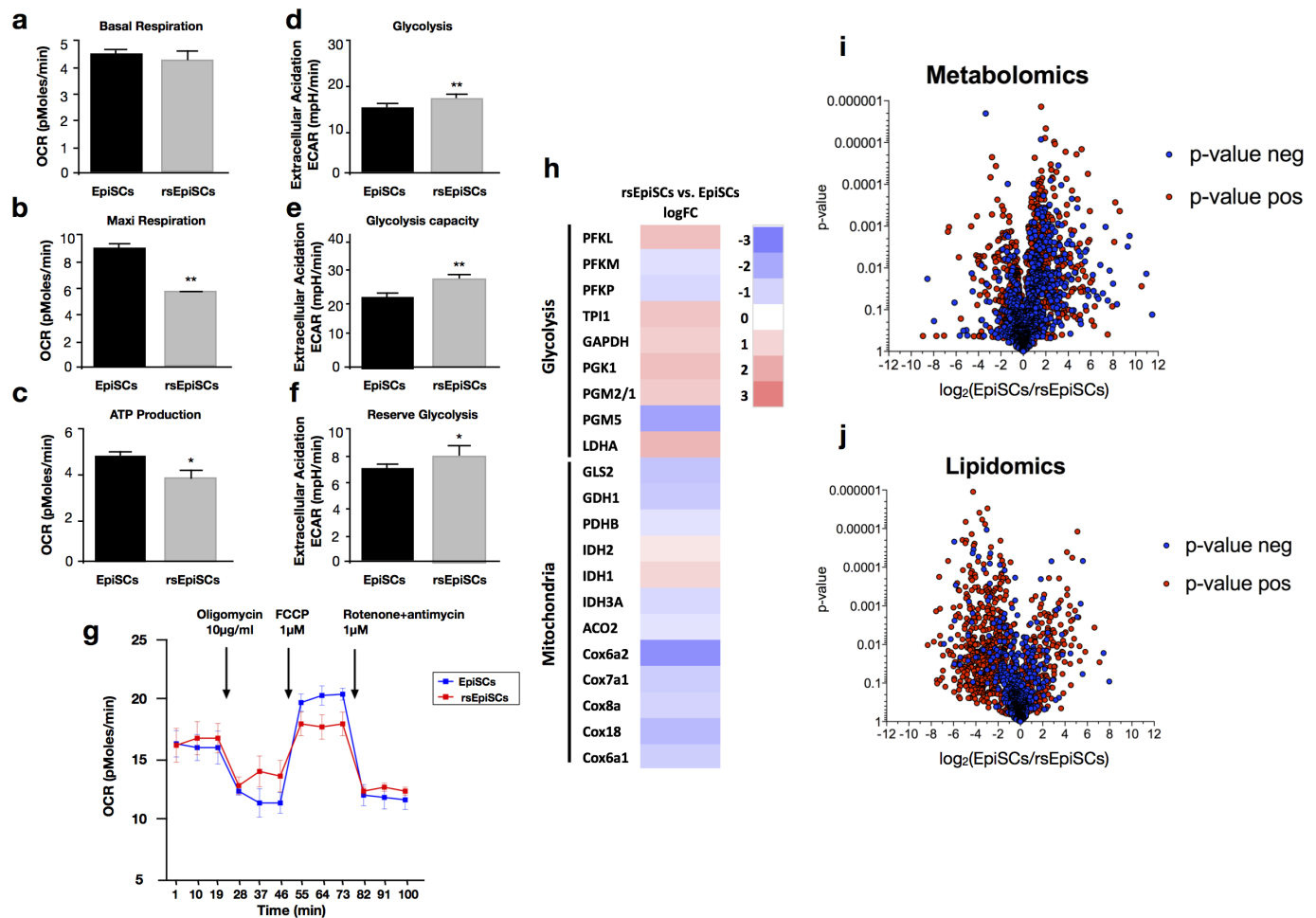
neuroectodermal (*Sox1* and *Pax6*) markers after indicated treatments for 4 days in culture. Error bars indicate s.d. ($n = 3$, biological replicates). **c**, Schematic representation of how different signalling pathways are involved in promoting or inhibiting self-renewal of rsEpiSCs.



Extended Data Figure 6 | Global transcriptomic and epigenomic analysis.

a, Hierarchical clustering of microarray gene expression data from ESCs, EpiSCs, rsEpiSCs and *in vivo* isolated E5.75 and E6.5 epiblasts. **b**, Two-way scatter plot of gene expression data from RNA-seq of EpiSCs and rsEpiSCs. Black lines indicate fourfold cut-off in expression level difference. Pearson correlation coefficient (r) between samples is shown at the upper right corner. **c**, Top five Gene Ontology (GO) terms enriched in the set of genes that are differentially expressed by at least fourfold (either up or down) between rsEpiSCs and EpiSCs. **d**, Average H3K27me3 signal at Polycomb target genes in rsEpiSCs (purple) and EpiSCs (green). **e**, Plots of DNA methylation and histone methylation (H3K4me3 and H3K27me3) signals around the transcription start sites of representative classes of genes. Examples given include pluripotent genes (*Sall4*, *Klf4* and *Lin28a*), neuronal-related genes (*Sox2*, *Gbx2* and *Sox1*) and cell-membrane-related genes (*Cldn6*, *Cldn3* and *Cdh1*). **f**, Global cytosine methylation at CG sites (mCG) levels of EpiSCs and rsEpiSCs (top left). The numbers of hyper- and hypo-DMRs discovered in rsEpiSCs (top right). The numbers of promoter-associated (transcription start site ± 2.5 kb), distal (>10 kb from transcription start site) and proximal (transcription start site ± 2.5 to 10 kb) rsEpiSCs hyper- and hypo-DMRs (bottom). **g**, Positive correlation between the amount of gene body non-CG DNA methylation and the level of gene expression. **h**, Gene Ontology biological process and molecular function term enrichment for genomic regions associated with rsEpiSCs hyper-DMRs.

i, PC1–PC2 plane from PCA analysis of transcriptome comparison between samples from this study (rsEpiSCs (circled with red line) and EpiSCs (circled with yellow line)) and a published data set⁶ (*in vivo* epiblast isolated from different developmental stages (CAV, cavity; PS, pre-streak; LMS, late mid-streak; LS, late streak; OB, no bud; EB, early bud; LB, late bud) and EpiSCs (circled with thick blue line)). The green arrow through the *in vivo* samples delineates a progressing 'developmental axis'. **j**, Hierarchical clustering of rsEpiSCs (red), EpiSCs (both from this study (yellow) and ref. 6 (blue)) and epiblasts of CAV, PS, LMS, LS, OB, EB and LB stages (green) using data from all annotated probes. **k**, Relative expression between rsEpiSCs and *in vivo* late-streak-stage epiblast (ref. 6) of genes characteristic of anterior mesendoderm (AME), anterior definitive endoderm (ADE), anterior primitive streak (APrS), whole primitive streak (PrS) and posterior primitive streak (PPrS). **l**, Primordial germ cell induction from Blimp1–YFP mESCs and rsEpiSCs. Left, before induction, both mESCs and rsEpiSCs were found negative for YFP; successful induction was observed with mESCs, as indicated by a positive YFP signal in cell aggregates, but not with rsEpiSCs. Right, PGC induction efficiency was compared between Blimp1–YFP mESCs and rsEpiSCs. Error bars indicate s.d. ($n = 3$, independent experiments); t -test, $**P < 0.01$, $*P < 0.05$. CAV, cavity; PS, pre-streak; ES, early streak; MS, mid-streak; LMS, late mid-streak; LS, late streak; OB/EB, no bud/early bud; LB, late bud.

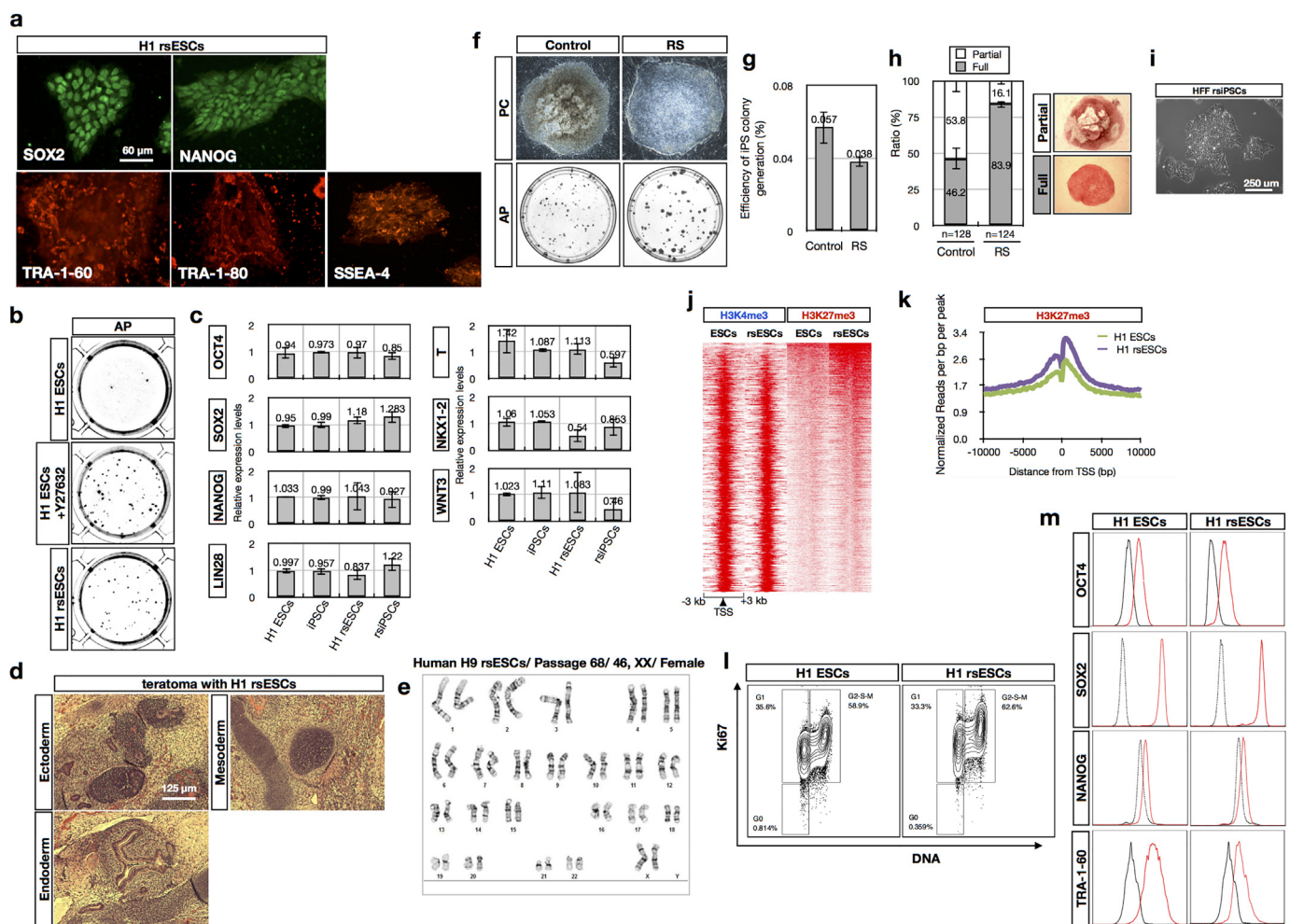


Extended Data Figure 7 | Metabolic profiling of EpiSCs and rsEpiSCs.

a, Basal respiration, **b**, **c**, Maximum respiration (**b**) and ATP production (**c**) were determined by calculating the average oxygen consumption rate (OCR) for each phase in EpiSCs and rsEpiSCs. **d**, **e**, **f**, Glycolysis (**d**), glycolytic capacity (**e**) and reserve glycolysis (**f**) were determined by calculating the average extracellular acidification rate (ECAR) for each phase in EpiSCs and rsEpiSCs. (Graph Pad Prism v5). **g**, Representative graph showing the oxygen consumption rate in response to oligomycin, FCCP and rotenone/antimycin of

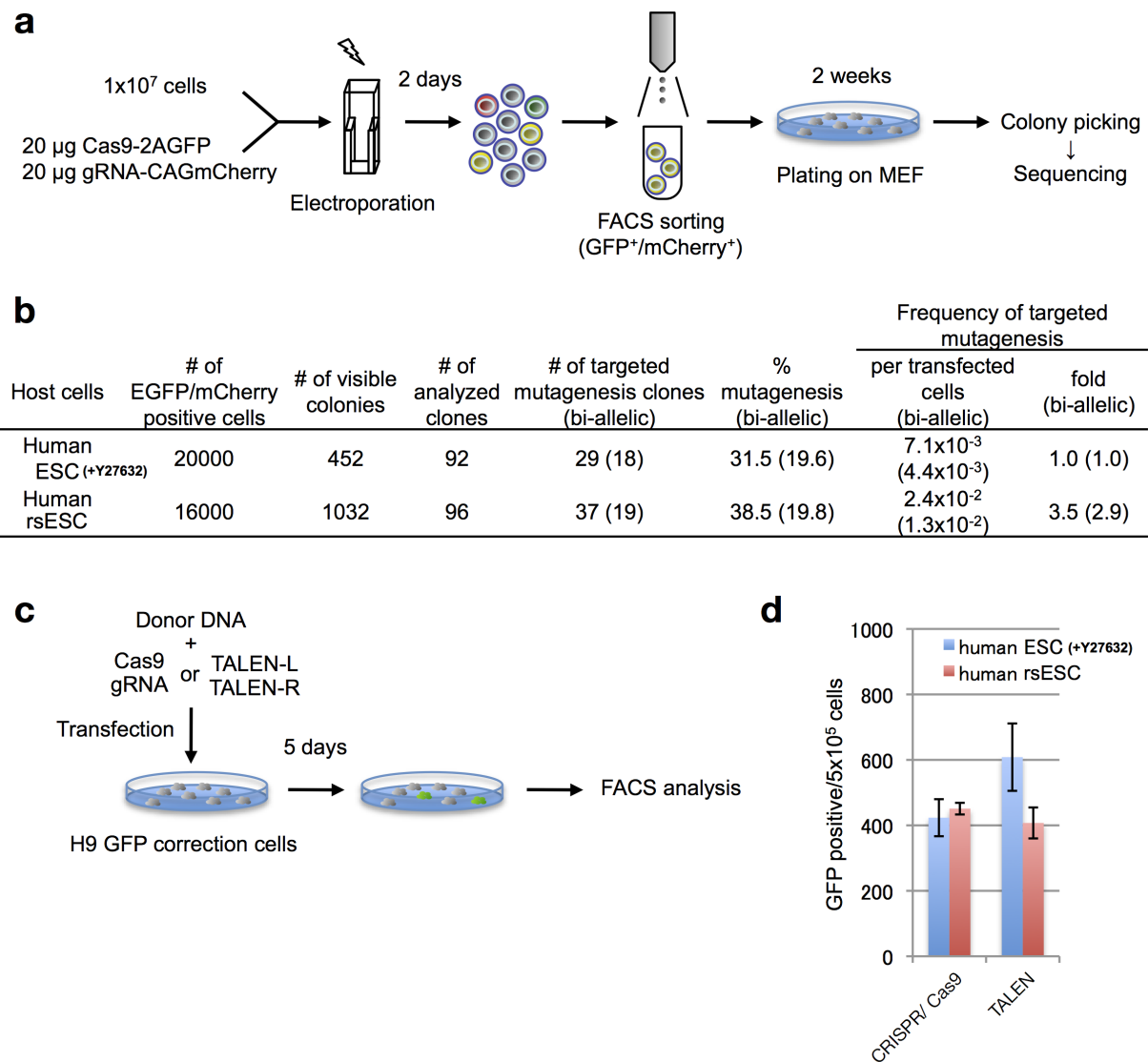
EpiSCs and rsEpiSCs ($n = 4$). **h**, Heat map of differentially expressed genes for mitochondrial complex COX and enzymes involved in glycolysis and the tricarboxylic acid cycle selected from the RNA-seq data set, $P < 0.05$.

i, **j**, Volcano plots of hydrophilic (metabolomics) and hydrophobic (lipidomics) metabolites show broad changes in metabolite levels between EpiSCs and rsEpiSCs. Error bars indicate s.d.; t -test, ** $P < 0.01$, * $P < 0.05$ (**a–g**, $n = 6$, technical replicates).



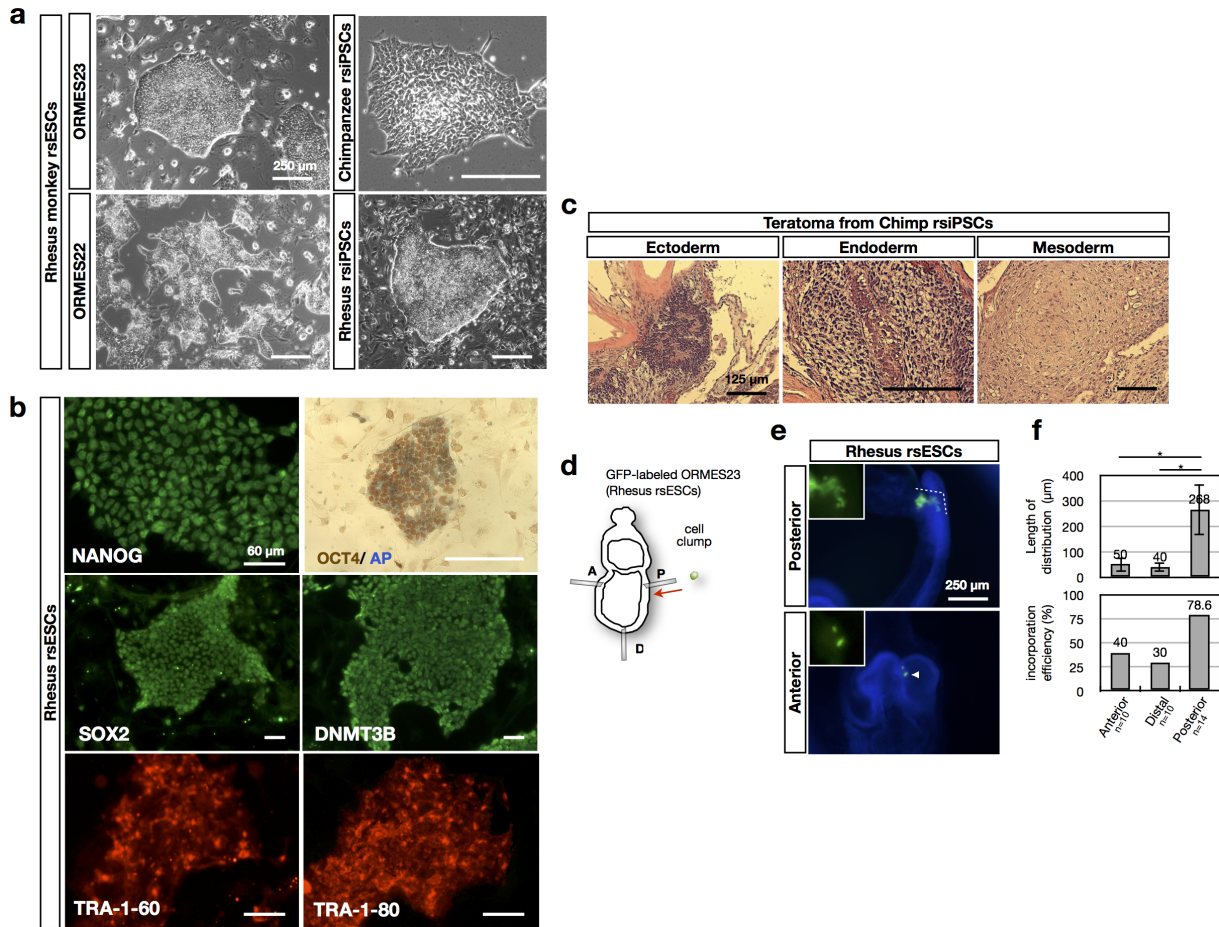
Extended Data Figure 8 | F/R1-based culture supports self-renewal of human ESCs as well as iPSC generation. **a**, Expression of pluripotency markers SOX2, NANOG, TRA-1-60, TRA-1-80 and SSEA-4 in human H1 rsESCs. **b**, Representative bright-field images showing colonies visualized by alkaline phosphatase (AP) staining after being plated at clonal density (1,000 cells per well) and cultured for 6 days. Y27632 was added at 10 μ M. **c**, Real-time quantitative PCR analysis of expression of pluripotency marker genes (*OCT4*, *SOX2*, *NANOG* and *LIN28*) and lineage marker genes (*T*, *NKX1-2* and *WNT3*) in H1 ESCs, human-foreskin-fibroblast-derived iPSCs, H1 rsESCs and human-foreskin-fibroblast-derived rs-iPSCs. Error bars indicate s.d. ($n = 3$, biological replicates). **d**, Haematoxylin and eosin staining images of teratomas generated from human H1 rsESCs show lineage differentiation towards three germ layers. **e**, Karyotype analysis of human H9 rsESCs indicates a normal diploid chromosome content. **f**, Representative bright-field images showing morphologies of putative iPSC colonies in conventional F/A-based human ESC culture and F/R1-based culture conditions (top). Alkaline phosphatase staining

at day 25 post-nucleofection indicates a larger colony size in F/R1-based culture (bottom). **g**, Efficiency of iPSC generation in conventional F/A-based human ESC culture conditions and F/R1-based culture conditions. Error bars indicate s.d. ($n = 3$, independent experiments). **h**, Quality of human iPSC-like colonies generated in F/A-based and F/R1-based culture conditions. Partial and full alkaline-phosphatase-positive iPSC-like colonies were counted separately using the criteria shown on the right. Error bars indicate s.d. ($n = 3$, independent experiments). **i**, Phase-contrast image showing morphology of human-foreskin-fibroblast-derived rs-iPSCs. **j**, Graphic representation of H3K4me3 and H3K27me3 ChIP-Seq signals near the transcription start site (TSS) for Polycomb target genes in H1 ESCs and H1 rsESCs. **k**, Average H3K27me3 signal at Polycomb target genes in H1 rsESCs (purple) and H1 ESCs (green). **l**, Cell-cycle profiles of H1 ESCs and H1 rsESCs were analysed by flow cytometry. **m**, Flow cytometry analysis of OCT4, SOX2, NANOG and TRA-1-60 protein expression in H1 ESCs and H1 rsESCs.



Extended Data Figure 9 | Genome editing in human rsESCs. **a**, Schematic representation of the targeted mutagenesis approach employed in human ESCs and rsESCs by CRISPR/Cas9. **b**, Targeted mutagenesis efficiencies at the *LRRK2* locus in human ESCs (treated with Y27632, 10 µM) and rsESCs. **c**, Schematic representation of the CRISPR/Cas9- or TALEN-mediated gene

correction approaches in human ESCs and rsESCs containing a mutated GFP gene. **d**, GFP correction efficiencies in human ESCs (treated with Y27632, 10 µM) and rsESCs by CRISPR/Cas9 or TALEN. The y axis shows the gene correction efficiency, which was calculated as GFP-positive cells per 5×10^5 cells. Error bars indicate s.d. ($n = 3$, independent experiments).



Extended Data Figure 10 | Non-human primate rPSCs. **a**, Phase-contrast images of colony morphologies of rhesus macaque rsESCs (ORMES22 and ORMES23), rhesus macaque rsiPSCs and chimpanzee rsiPSCs.

b, Immunofluorescence images of NANOG, SOX2, DNMT3b, TRA-1-60 and TRA-1-80 protein expression in ORMES23 rsESCs. ORMES23 rsESCs were also stained for alkaline phosphatase (AP) activity and OCT4 immunohistochemistry (top right). **c**, Haematoxylin and eosin staining images of teratomas generated by chimpanzee rsiPSCs show lineage differentiation towards three germ layers. **d**, Schematic representation of epiblast grafting experiment with GFP-labelled ORMES23 rsESCs (Please refer to Supplementary Fig. 1 for details). A, anterior; P, posterior; D, distal.

e, Fluorescence images of grafted embryos after *in vitro* culture. GFP-labelled

ORMES23 rsESCs were grafted to posterior, distal and anterior regions of epiblasts of isolated non-intact and non-viable E7.5 mouse embryos and cultured *in vitro* for 36 h before fixation and visualization by an inverted fluorescence microscope. Arrowhead indicates a cell clump failed to distribute. Dashed line indicates dispersed cells in the posterior region of grafted embryo. Blue, DAPI. Insets show higher-magnification images of GFP-labelled cells. **f**, Top, quantification of the extent of cell spreading of GFP-labelled ORMES23 rsESCs after being grafted to different regions of E7.5 mouse epiblasts. Bottom, incorporation efficiency of grafted GFP-labelled ORMES23 rsESCs in mouse E7.5 epiblasts. Error bars indicate s.d. (*n*, indicated on the graph, independent experiments); *t*-test, **P* < 0.05.

Neurotransmitter and psychostimulant recognition by the dopamine transporter

Kevin H. Wang^{1†*}, Aravind Penmatsa^{1†*} & Eric Gouaux^{1,2}

Na⁺/Cl[−]-coupled biogenic amine transporters are the primary targets of therapeutic and abused drugs, ranging from antidepressants to the psychostimulants cocaine and amphetamines, and to their cognate substrates. Here we determine X-ray crystal structures of the *Drosophila melanogaster* dopamine transporter (dDAT) bound to its substrate dopamine, a substrate analogue 3,4-dichlorophenethylamine, the psychostimulants D-amphetamine and methamphetamine, or to cocaine and cocaine analogues. All ligands bind to the central binding site, located approximately halfway across the membrane bilayer, in close proximity to bound sodium and chloride ions. The central binding site recognizes three chemically distinct classes of ligands via conformational changes that accommodate varying sizes and shapes, thus illustrating molecular principles that distinguish substrates from inhibitors in biogenic amine transporters.

Signals by the biogenic amine neurotransmitters—dopamine (DA)¹, serotonin² and noradrenaline³—at chemical synapses are terminated by the cognate neurotransmitter sodium symporters (NSSs)^{4–7}. Biogenic amines play profound roles in the development and function of the nervous system, as well as in animal behaviour and activity; thus NSSs are central to normal neurophysiology and are the targets of a spectrum of therapeutic and illicit agents, from antidepressants and anti-anxiety medications to cocaine and amphetamines⁸. Experimental and computational studies have shown that the DA, serotonin (SERT) and noradrenaline (NET) transporters harbour a conserved structural fold^{9,10}, first seen in the structure of LeuT¹¹. Owing to variations in amino acid sequences¹², however, the biogenic amine transporters possess distinct yet overlapping pharmacological ‘fingerprints’¹³.

The dopamine transporter (DAT)¹⁴ removes DA from synaptic and perisynaptic spaces, thus extinguishing its action at G-protein coupled DA receptors. To drive the vectorial ‘uphill’ movement of extracellular DA into presynaptic cells, DAT couples substrate transport to pre-existing sodium and chloride transmembrane gradients. Congruent with the multifaceted roles of DA in the nervous system, perturbation of dopaminergic signalling by disruption of native DAT function has profound consequences^{15–17}. On the one hand, the amphetamines, potent and widely abused psychostimulants, are DAT substrates that enhance synaptic levels of DA both by competing with DA transport by DAT and by inducing the release of DA from synaptic vesicles into the cytoplasm, from where DA is then effluxed through DAT into the synaptic space^{18–24}. On the other hand, the *Erythroxylum coca* leaf-derived alkaloid, cocaine, as well as synthetic cocaine derivatives are competitive inhibitors of DAT and enhance extracellular DA concentrations by locking the transporter in a transport inactive conformation^{14,25–27}. Widely prescribed antidepressants specifically inhibit serotonin and noradrenaline uptake and typically have weaker affinities towards DAT^{28,29}.

Mutagenesis, chemical modification, binding and transport studies have implicated the central or S1 binding site in DAT, akin to the leucine and tryptophan site in LeuT, as the binding site occupied by DA, amphetamines, cocaine and antidepressants^{25,26,30}. Moreover, the X-ray structure of a transport-inactive *Drosophila melanogaster* DAT

(dDAT) in complex with nortriptyline shows the antidepressant bound at the central site^{9,31}. Nevertheless, none of these studies have visualized the binding of DA, amphetamine or cocaine to an active DAT, nor have they illuminated distinctions in ligand pose and transporter conformation between substrates and inhibitors. Here we present X-ray structures of dDAT with substrates DA, methamphetamine or D-amphetamine, with the DA analogue 3,4-dichlorophenethylamine (DCP), and with cocaine or cocaine analogues.

Resurrection of transport activity

The previously reported structure of the dDAT–nortriptyline complex exploited a transport-inactive variant with five thermostabilizing mutations (dDAT_{cryst})⁹. We recovered transport function yet retained favourable crystallization properties by reverting three thermostabilizing mutations (V275A, V311A and G538L) to their wild-type identities and by shifting the deletion of extracellular loop 2 (EL2; Extended Data Fig. 1). This minimal functional construct, dDAT_{mfc}, has a melting temperature of 48 °C³², exhibits DA transport with a K_M of $8.2 \pm 2.3 \mu\text{M}$ and V_{max} of $2.4 \pm 0.2 \text{ pmol min}^{-1}$ per 10^6 cells, compared to wild-type dDAT (dDAT_{wt}) with a K_M of $2.1 \pm 0.7 \mu\text{M}$ and V_{max} of $4.5 \pm 0.4 \text{ pmol min}^{-1}$ per 10^6 cells (s.e.m., Fig. 1a). The dDAT_{mfc} construct binds nisoxetine with a K_d of 36 nM compared to a K_i of 5.6 nM for wild-type dDAT³¹ (Extended Data Fig. 2).

The central binding site in DAT, NET and SERT can be divided to subsites A, B and C^{29,33}. Subsites A and C are well conserved in dDAT versus human DAT (hDAT), whereas subsite B, a pocket sculpted by TMs (transmembrane helices) 3 and 8, differs from hDAT in that residues lining this pocket in dDAT are Asp121 and Ser426 (Extended Data Fig. 3). We introduced mutations D121G (TM3) and S426M (TM8) into the dDAT_{cryst} and dDAT_{mfc} constructs to mimic hDAT subsite B³³. These mutations enhanced the affinities for nisoxetine, β -CFT (2 β -carbomethoxy-3 β -(4-fluorophenyl)tropane) and DCP (Extended Data Figs 2, 4). Although constructs harbouring subsite B substitutions improved crystallization propensity, transport activity was extinguished (Extended Data Fig. 3c). Nevertheless, structures bearing these mutations were solved in complexes with cocaine, β -CFT, RTI-55(2 β -carbomethoxy-3 β -(4-iodophenyl)tropane)

¹Vollum Institute, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. ²Howard Hughes Medical Institute, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. [†]Present addresses: Amgen, Division of Molecular Structure and Characterization, Cambridge, Massachusetts 02142, USA (K.H.W.) and Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India (A.P.).

*These authors contributed equally to this work.

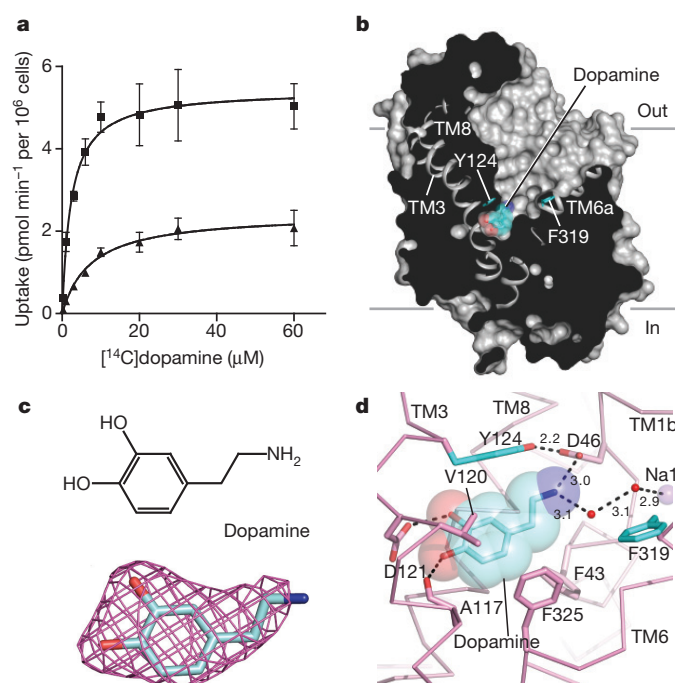


Figure 1 | Dopamine occupies central binding site. **a**, Michaelis-Menten plots of specific DA uptake by dDAT_{wt} (squares) and dDAT_{mfc} (triangles). Graph depicts one representative trial of two independent experiments. Error bars represent s.d. values for technical replicates measured in triplicate. **b**, Surface representation of the dDAT_{mfc}-DA complex viewed parallel to the membrane, with DA displayed as cyan spheres. Residues Y124 and F319 are shown as cyan sticks on the left and right sides of DA, respectively. **c**, Chemical structure and $F_o - F_c$ density for DA (3.0σ). **d**, Close-up view of DA in the binding pocket with hydrogen bonds shown as dashed lines. Sodium ions and water are shown as purple and red spheres, respectively.

or DCP (Supplementary Table 1). In the cocaine, RTI-55 and DCP complexes, superposition of structures with subsite B mutations onto structures of dDAT_{mfc} complexes did not reveal prominent structural changes in the binding pocket or deviations in the positions of bound ligand (Extended Data Table 1).

Dopamine bound to central binding site

The structure of dDAT_{mfc} bound to DA displays an outward-open conformation (Fig. 1b) where DA is situated in the central binding site, surrounded by TMs 1, 3, 6 and 8. The amine group points towards subsite A and interacts with the carboxylate of Asp46 at a distance of 3 Å. The catechol group occupies a subsite B cavity sculpted by Ala117, Val120, Asp121, Tyr124, Ser422 and Phe325 (Fig. 1c, d), residues predicted to interact with DA using homology models of DAT based on the substrate-bound occluded state of LeuT, analysis of uptake kinetics, and cysteine labelling studies^{10,25,26,34}. Despite a lack of steric interference from DA, Phe319, which is equivalent to Phe253 in LeuT (in which it occludes solvent access to the LeuT binding pocket), remains splayed away from DA, in an orientation seen in the nortriptyline-bound structure^{9,11}.

The location and interactions of DA with dDAT recall predictions made through homology models of hDAT, although discrepancies between the cocrystal structure and the homology models are also evident. Site-directed mutagenesis and molecular dynamics simulations pointed to the crucial role of Asp46 (Asp79 in hDAT) in the recognition of DA^{10,25,34,35}, a residue conserved amongst biogenic amine neurotransmitters. GAT (γ-amino butyric acid transporter), GlyT (glycine transporter), and LeuT contain glycine at the equivalent position owing to the presence of a compensatory carboxylate group in the substrates¹¹. One notable difference between the dDAT_{mfc}-DA structure and previous hDAT models, however, is a rotation in the χ₁

torsion angle of ~130° in the side chain of Asp46 to maintain a 3 Å distance to the amine group of DA. This rotation severs the indirect coordination between Asp46 and the sodium ion at site 1 as observed in the nortriptyline-dDAT_{cryst} complex⁹. Unlike previous simulation studies where DA binds in a dehydrated pocket³⁴, we observed non-protein electron density 3.1 Å from the amine group of DA, into which a water molecule was modelled. This water molecule forms a hydrogen bond with the water molecule that coordinates the sodium ion at site 1, resulting in a molecular network that links DA to the ion-binding sites. In the case of LeuT, substrate interaction with sodium site 1 is direct, with the carboxylate of leucine coordinating the sodium ion¹¹.

The catechol ring of DA interacts with TMs 3 and 8 by hydrogen bonds with the carboxylate group of Asp121 (Fig. 1d). Modelling studies predicted that DA occupies multiple poses within the binding pocket with the *meta*-hydroxyl either interacting with residues in TM8 equivalent to Ser421 or to Ser422 (dDAT numbering)³⁴. By contrast, in the structure of dDAT_{mfc}-DA, the *para*-hydroxyl group interacts with both the carbonyl oxygen of Ala117 and the carboxylate of Asp121 at distances of 2.8 and 3.1 Å, respectively, whereas the *meta*-hydroxyl group interacts with the side chain of Asp121 at a distance of 2.7 Å and faces Ser422 in TM8 at a distance of 3.8 Å.

The residues poised to interact with the catechol ring vary across DAT orthologues with invertebrate DAT orthologues retaining alanine and aspartate at positions equivalent to residues 117 and 121, respectively, whereas in most mammalian DATs residue Asp121 is replaced by glycine and Ala117 by serine⁹, the latter of which could act as a surrogate hydrogen bond partner for the catechol group. With hNET, the equivalent residues at 117 and 121 are Ala and Gly, raising the possibility of the catechol group of noradrenaline interacting with the hydroxyl of Ser420 in hNET (TM8; equivalent to Ser422 of dDAT and Ala423 in hDAT). We propose that these compensatory variations within subsite B dictate catecholamine recognition common to both DAT and NET. To explain why NET binds noradrenaline and DA with nearly equal apparent affinity yet DAT prefers DA³¹, we must invoke longer-range, indirect interactions, perhaps involving subsite B and the non-helical TM6a-6b 'linker' because the 'Phe-box' surrounding the β-carbon position is conserved between NET and DAT.

Recognition of D-amphetamine and methamphetamine

To understand how amphetamines are transported by DAT despite lacking the hydroxyl groups of the catecholamines, we characterized the interactions between amphetamines and dDAT by binding assays and crystallographic studies. (+)-methamphetamine displaces [³H]nisoxetine binding to dDAT_{mfc} with a K_i value of 31 μM, whereas D-amphetamine, which lacks the *N*-methyl group present in methamphetamine, has a K_i of 86 μM (Fig. 2a). D-amphetamine is 10 to 100-fold weaker in its ability to inhibit DA transport in the dDAT compared to its mammalian counterparts³¹. The weaker affinities of dDAT for amphetamines compared to mammalian DATs may be due in part to differences in residues of subsite B. Indeed, the presence of Asp121 and Ser426 in invertebrate DATs creates a polar environment that does not complement the non polar benzyl groups of amphetamines (Fig. 1d). In mammalian DATs both methamphetamine (K_i = 0.5 μM) and D-amphetamine (K_i = 0.6 μM) are nearly as effective as cocaine (0.2 μM) at inhibition of DA uptake³⁶. Although the maximal rate of transport (V_{max}) for D-amphetamine in hDAT is fivefold lower than for DA, the K_M values range from 0.8 to 2 μM²², consistent with the notion that mammalian subsite B is more complementary towards the binding of amphetamines than subsite B in invertebrate DATs.

The structures of (+)-methamphetamine-dDAT_{mfc} and D-amphetamine-dDAT_{mfc} displayed outward open conformations with electron densities for the drugs found in the central binding site (Extended Data Fig. 5; Fig. 2b). The amine groups of methamphetamine

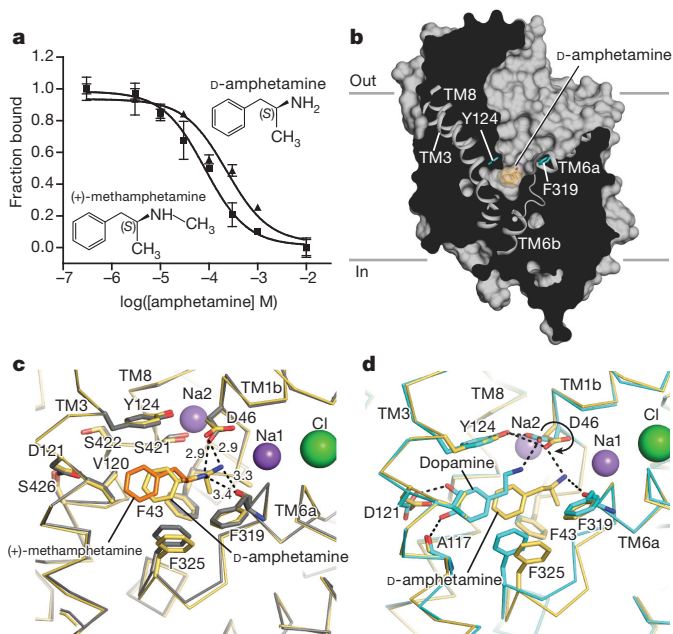


Figure 2 | Amphetamines bind to central site. **a**, Displacement of bound [^3H]nisoxetine by methamphetamine ($K_i = 31 \mu\text{M}$; squares) and D-amphetamine ($K_i = 86 \mu\text{M}$; triangles). Graph depicts one representative trial of two independent experiments. Error bars represent s.e.m. values for technical replicates measured in triplicate. **b**, Surface representation of D-amphetamine-dDAT_{mfc} complex viewed parallel to the membrane with ligands shown as a light orange sphere. Residues Y124 and F319 are shown as cyan sticks on the left and right sides of DA, respectively. **c**, **d**, Superposition of binding pockets of the D-amphetamine-dDAT_{mfc} structure in pale orange with binding pockets of methamphetamine-dDAT_{mfc} (**c**, drug in orange, backbone in grey) and DA-dDAT_{mfc} in teal (**d**). Hydrogen bond interactions are represented as dashed lines. Asp46 undergoes a χ_1 torsion angle shift from -168° in DA-bound state to $+62^\circ$ in the D-amphetamine-dDAT_{mfc}.

and D-amphetamine lie closer to Asp46 at subsite A with hydrogen bonding distances of 2.9 Å, and the main chain carbonyl of Phe319 is positioned nearby at 3.3 Å (Fig. 2c). The amine group of D-amphetamine interacts with Asp46, which does not undergo the rotameric shift as seen in the DA-bound structure because D-amphetamine is situated in the centre of the pocket and not displaced by 2.8 Å towards TMs 3 and 8 as seen with DA (Fig. 2d). By contrast with earlier findings²⁵, we do not observe a disruption of the hydrogen bond between Asp46 and Tyr124 despite Asp46 clearly forming a hydrogen bond with the primary amine of D-amphetamine. Phe325 retains edge-to-face aromatic interactions with the phenyl group of amphetamines by way of a contraction of the TM6a–6b linker in comparison to the DA-bound state. Amphetamines adopt poses in the central binding that allow for interactions between their amino and aromatic groups and transporter subsites, thus explaining how the sterically smaller amphetamines compete with DA and act as substrates despite the absence of catechol-like hydroxyl groups.

Dopamine analogue stabilizes partially occluded state

DA is prone to oxidation and thus we sought a stable analogue for crystallographic and biochemical studies. Multiple high-affinity biogenic amine transporter inhibitors harbour halogen groups on the aromatic rings predicted to occupy subsite B^{33,37} and thus we screened halogenated phenethylamine derivatives for binding to dDAT (Extended Data Fig. 4d). We discovered that 3,4-dichlorophenethylamine (DCP) possessed the greatest affinity and, because it is approximately isosteric to DA, was selected for further study (Fig. 3a).

One DCP molecule is lodged in the central binding pocket, with the amine group forming a hydrogen bond with Asp46, and the

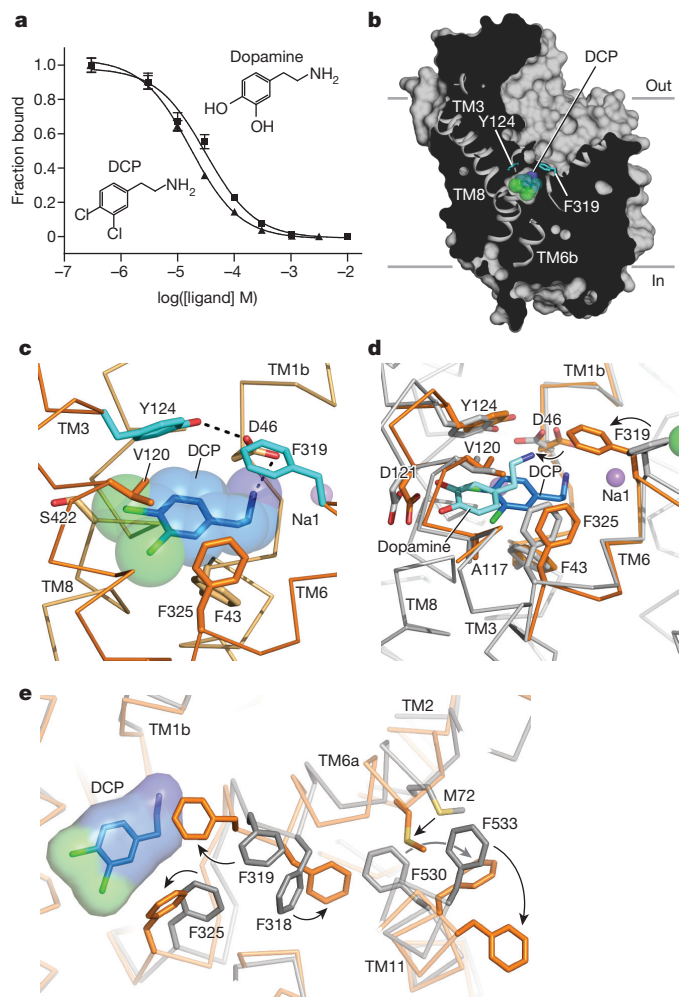


Figure 3 | DCP induces partial occlusion of central site. **a**, DCP and DA inhibit binding of [^3H]nisoxetine to dDAT_{mfc} with K_i constants of 4.5 and 8.3 μM , respectively. Graph depicts one representative trial of two independent experiments, and data points indicate average values of technical replicates measured in triplicate for DA (squares) and DCP (triangles). Error bars represent s.d. values for individual data points measured in triplicate. **b**, Surface representation of the DCP-dDAT_{mfc} complex viewed parallel to the membrane with DCP shown as blue spheres and chloride in green. Residues Y124 and F319 are shown as cyan sticks on the left and right sides of DCP, respectively. **c**, Close-up view of DCP in the binding pocket showing the hydrogen bond between the primary amine of DCP and D46 with a distance of 3.2 Å. Sodium ions are shown as purple spheres. **d**, Superposition of DA- and DCP-bound dDAT_{mfc} structures reveals a 3.1 Å displacement in ligand position. The DA-dDAT_{mfc} and DCP-dDAT_{mfc} structures are coloured grey and orange, respectively. **e**, Conformational changes of phenylalanine side chain positions in the partially occluded state of DCP-dDAT_{mfc} compared to the outward open state of nortriptyline-dDAT_{cryst} coloured orange and grey, respectively.

dichlorophenyl ring bordered by Val120 and Phe325 in subsites B and C, respectively (Fig. 3b–d, Extended Data Fig. 6). The position of DCP in the pocket is closer to that of D-amphetamine in dDAT and of substrate leucine in LeuT¹¹, whereas the position of DA is shifted towards TMs 3 and 8, probably owing to hydrogen bonding between the catechol hydroxyl groups and Asp121. As a result, the position of Asp46 in the DCP-bound structure is superimposable with the positions seen in the amphetamine-bound structures. Interestingly, the side chain of Phe319 rotates to occlude the binding pocket, a conformational change not seen in any of the inhibitor- or DA-bound structures (Fig. 3c, d). This rotamer of Phe319 prevents solvent access to the pocket, leaving an aperture only ~ 1 Å wide on the extracellular side of DCP. The equivalent residue in LeuT, Phe253, adopts the same

orientation in the occluded substrate-bound form, supporting the notion that the rearrangements seen in the DCP-dDAT_{mfc} structure are on the pathway to a LeuT-like occluded state of dDAT¹¹. Unlike the DA-bound state, there is no evidence of a water molecule associated with DCP in the structures, suggesting that formation of an occluded state is associated with dehydration of the binding pocket (Fig. 3c, d), similar to LeuT.

The partially occluded binding pocket of the DCP-dDAT_{mfc} structure primarily results from rotations of TM1b and 6a 'into' the binding pocket, towards scaffold helices 3 and 8 and around axes centred near the non helical regions of TMs 1b and 6a (Extended Data Fig. 6a, c). In LeuT, studies of the transporter in solution and in the crystal show that TMs 1b and 6a, along with EL4, undergo conformational changes to close the extracellular gate^{11,38,39}. Comparisons with the outward-open nortriptyline-bound state of dDAT⁹ and the occluded state of LeuT indicate that, although the DCP ligand is nearly inaccessible to the extracellular solution, the inward rotations of TMs 1b (5.6°) and 6a (~7°) are less pronounced in dDAT than in LeuT¹¹ (Extended Data Fig. 6f, g). Nevertheless, these helical rotations position the side chain of Phe319 over the extracellular face of the binding pocket, and are associated with a series of phenylalanine side chain reorientations (Fig. 3e). TM11 undergoes an outward movement of 6° to accommodate these side chain shifts, and an inward movement of 5° is observed in TM2 (Extended Data Fig. 6b, d). Interestingly, TMs 2, 7 and 11 at the inner leaflet of the plasma membrane form a second cholesterol binding site (site 2) where a density for cholesterol hemisuccinate (CHS) was observed in all the reported structures. Indeed, CHS enhances dDAT inhibitor affinity and we speculate that, in native membranes, cholesterol binds to sites 1 and 2, perhaps stabilizing DAT in an outward-open state (Extended Data Fig. 7)^{39–41}.

The occluded state is a Michaelis–Menten-like transport intermediate for LeuT and related transporters such as BetP, Mhp1 and MhsT^{42–44} when bound to substrate. Disparities in conformations observed between LeuT and dDAT in the presence of substrate, however, may reflect bona fide differences that are dependent on the relative stabilities of distinct substrate-bound states. Nevertheless, we cannot exclude crystal lattice effects, Fab binding, lipid, or solutes present in the crystallization solutions that may favour the outward-open conformation observed for the dDAT–substrate complexes reported here.

Cocaine binds in the central site

The structure of the dDAT_{mfc}–cocaine complex exhibits an outward-open conformation (Fig. 4a) with cocaine bound to the central pocket at a site overlapping the nortriptyline site⁴⁵, adjacent to the Na1 and Na2 sodium ions and the chloride ion. The tertiary amino group of cocaine forms a salt bridge with Asp46 (TM1b) (Fig. 4b). We observe that the TM6a–6b linker contracts, allowing Phe325 to form edge-to-face aromatic interactions with the benzyl ring of cocaine (Fig. 4c).

The negative electrostatic potential of subsite B in dDAT compared to mammalian DATs probably underlies the reduced affinity of dDAT for cocaine. Altering this charged pocket to mimic mammalian DATs with the mutations D121G and S426M yielded a dDAT construct with enhanced binding affinities for cocaine and β -CFT, but not for RTI-55 (Fig. 4d, Extended Data Fig. 4a, b). We speculate that the carboxylate group of Asp121 does not form favourable interactions with the 4-fluorophenyl and 4-iodophenyl groups of β -CFT and RTI-55, respectively, perhaps accounting for some of the discrepancies in relative binding affinities between hDAT and dDAT.

Ligand docking studies using a homology model of DAT, in combination with biochemical binding assays and ligand-dependent disulphide trapping experiments, have probed the orientation of cocaine in the central binding site^{25,26,30}. These studies predicted that the fluorophenyl moiety of β -CFT forms a hydrogen bond with the side chain equivalent to Asn125 in dDAT. Furthermore, the methyl ester of cocaine was thought to displace the side chain of Tyr124 to

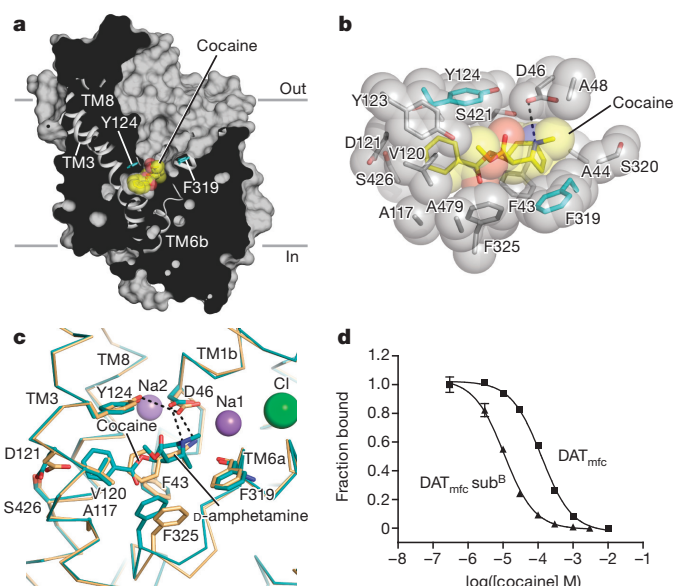


Figure 4 | Multivalent binding of cocaine. **a**, Surface representation of the structure of cocaine–dDAT_{mfc} viewed parallel to the membrane, with cocaine displayed as yellow spheres. Residues Y124 and F319 are shown as cyan sticks on the left and right sides of cocaine, respectively. **b**, Close-up view of cocaine in the binding pocket with residues interfacing with cocaine shown as spheres. The tertiary amine of cocaine is 3.4 Å from the carboxylate of D46. **c**, Superposition of the binding pocket of the D-amphetamine–dDAT_{mfc} structure in pale orange with the binding pocket of the cocaine–dDAT_{mfc} structure in teal. **d**, Displacement of [³H]nisoxetine by cocaine for dDAT_{mfc} (squares) and dDAT_{mfc} sub^B (triangles) constructs with inhibition constant (K_i) values of $33 \pm 3 \mu\text{M}$ and $3 \pm 0.3 \mu\text{M}$, respectively. Graph depicts one representative trial of two independent experiments, and data points indicate average values of technical replicates measured in triplicate.

abrogate the hydrogen bond between Tyr124 and Asp46 as a mechanism of inhibiting transporter function. The dDAT_{mfc}–cocaine structure contradicts the finding that cocaine disrupts the Asp46–Tyr124 interaction and does not place the side chain of Asn125 in a location where it could interact with the fluorophenyl group of β -CFT.

To validate the cocaine-bound structure and conclusively identify residues that interact with tropane-based ligands, structures of dDAT were solved in the presence of the cocaine analogues β -CFT and RTI-55. Anomalous scattering by the iodide of RTI-55 corroborated the location and placement of the aromatic moiety of cocaine proximal to TMs 3 and 8 (Extended Data Fig. 8a). The $F_o - F_c$ 'omit' electron density maps for cocaine, β -CFT, and RTI-55 are consistent with the methyl ester group protruding into the base of the extracellular vestibule without disrupting the Asp46–Tyr124 interaction (Extended Data Figs 5, 8b). The position of β -CFT places the fluoro group 6 Å from the amide nitrogen of Asn125, indicating that Asn125 does not directly participate in binding of β -CFT. Superpositions of the cocaine–dDAT_{mfc}, β -CFT–dDAT_{cryst}, and RTI-55–dDAT_{mfc} structures exhibited overall mean root-mean-square deviation (r.m.s.d.) values below 0.7 Å, and residues that make close contacts with the ligands overlap nearly entirely, indicating that the halide-substituted phenyl groups of β -CFT and RTI-55 do not markedly affect the architecture of the binding pocket (Extended Data Table 1).

Residues in the binding pocket that interact with cocaine are shared by β -CFT and RTI-55, with slight deviations in subsite B owing to the presence of halide substituents on these latter two inhibitors. At the distal end of the ligand, the halophenyl rings of β -CFT and RTI-55 or the benzoate of cocaine form van der Waals interactions primarily with Ala117, Val120, Tyr124, Phe325 and Ser422, all of which

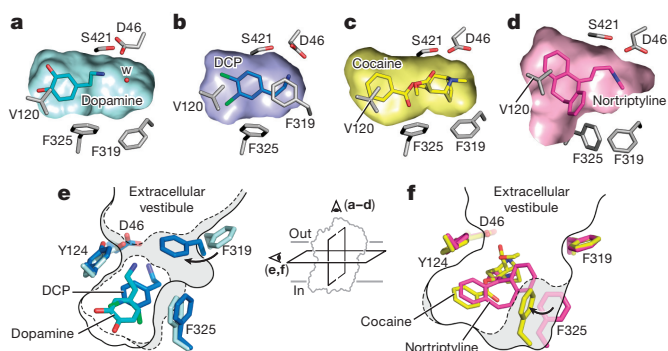


Figure 5 | Plasticity confers versatile recognition. **a–d**, Transverse sections of the binding pocket of dDAT_{mfc} are shown as surface representations in the DA- (**a**), DCP- (**b**), cocaine- (**c**) and nortriptyline-bound (**d**) structures. A water molecule (W) is observed in the vicinity of DA in **a**. **e, f**, Schematic representing plasticity within the substrate and drug binding pockets in occluded state of DCP–dDAT_{mfc} (broken lines) with DA–dDAT_{mfc} (solid line) (**e**) and cocaine–dDAT_{mfc} (broken lines) with nortriptyline–dDAT_{cryst} (**f**, grey, PDB ID 4M48) (solid line). Inset represents schematic to show field of view for **a–d** and **e, f**.

were previously predicted to interact with these tropane-based inhibitors^{25,46,47} (Fig. 4b). A comparison with the antidepressant nortriptyline–dDAT_{cryst} structure⁹ reveals that the smaller benzoate group of cocaine is accommodated with a shift of the TM6a–6b linker and a rotation of Phe325 into the binding pocket, which was not previously predicted. The tropane rings of cocaine, β -CFT and RTI-55 are bordered by Phe43, Ala44, Asp46, Ala48, Phe319 and Ser421. The side chain of Phe319 maintains a similar orientation as seen in the dDAT–nortriptyline structure owing to the bulky tropane ring and the methyl ester group present in all three tropane inhibitors. Overall, the docking and photolabelling studies predicted the location and orientation of cocaine and its analogues to overlap with the DA-binding site of DAT^{25,26,30}. Our structures validate these findings yet indicate distinctions that have implications for the mechanism of inhibition by tropane ligands. Taken together, structures of dDAT in complex with tropane ligands suggest that inhibition is achieved by combining a free amine with bulky tropane and aromatic moieties to limit conformational movements of the transporter.

Plasticity of the substrate binding pocket

The inhibitor- and substrate-bound structures of dDAT indicate that the binding pocket of dDAT accommodates ligands of varying sizes largely by adjusting the orientation of the discontinuous region of TM6 and the side chains of Phe319 and Phe325 (Fig. 5; Extended Data Table 2a). Ligand recognition by dDAT is bipartite and requires an amine group that interacts with the carbonyl oxygens of Phe43 and Phe319 or the carboxylate of Asp46 in subsite A, in combination with an aromatic group that is stabilized by van der Waals interactions with residues lining a hydrophobic cleft formed by TMs 3, 6, and 8 of subsite B. The substrates DA, D-amphetamine, and methamphetamine each contain one phenyl ring linked to an ethylamine chain, requiring Phe325 to rotate inward to contract the size of the pocket and maintain edge-to-face aromatic interactions with these ligands (Extended Data Fig. 8c, Fig. 5a, e). In the DCP–dDAT_{mfc} structure, Phe319 also rotates inward to cover the ligand, and similar to LeuT, this reorientation of Phe319 is required for the formation of the occluded state (Extended Data Fig. 8d, Fig. 5b, e). In order for Phe319 to cover DCP in the pocket, rotation of Phe319 is accompanied by the inward tilting of TM6a to bring the hinge region of TM6 towards the ligand (Extended Data Fig. 6). Accommodation of the single phenyl rings in tropane inhibitors requires Phe325 to assume a position similar to that seen in the DA-bound structure to provide edge-to-face aromatic interactions (Fig. 5c, f). To enlarge the binding pocket for multiple aromatic rings, the side chains of Phe319 and

Phe325 are splayed outward as seen in the antidepressant-bound structures (Fig. 5d, f).

These structures provide a molecular explanation for the distinction between substrates and inhibitors of biogenic amine transporters. Substrates such as DA and amphetamines contain amine and aromatic functional groups at opposing ends of the molecule that interact with both the extracellular gate of TMs 1b and 6a and the scaffold TMs 3 and 8. In contrast, inhibitors exploit the flexibility of the binding pocket to bind to the outward open transporter with high affinity, acting like wedges to lock the transporter in an outward-open conformation. Tropane ligands achieve inhibition by inserting benzyl or halo-phenyl groups into the cavity of subsite B, with the tropane ring oriented to sterically hinder the movement of the extracellular gate. Antidepressants, like nortriptyline, differ from tropane inhibitors by coupling bulky aromatic moieties with an amine group to block conformational flexibility in the transporter.

Conclusions

Structures of dDAT in complex with both substrates and inhibitors emphasize the role of subsites in the pocket of dDAT in defining ligand specificity, a concept that can be expanded to understand variation in pharmacological profiles between biogenic amine transporters³³. The DA-bound structure suggests that interactions between the catechol ring and subsite B, together with hydrogen bonding between the amine group of DA and Asp46, drive closure of the extracellular gates TMs 1b and 6a to form the occluded state. D-amphetamine and methamphetamine are bound in a manner distinct from DA in the pocket, and the absence of hydroxyl groups in amphetamines indicates that hydrophobic interactions must be sufficient for amphetamines to interact with subsite B residues and bridge the scaffold TMs 3 and 8 with the extracellular gating helices. The conformations of the DA, amphetamine and DCP-bound complexes probably represent snapshots following substrate and ion binding but before full closure of the extracellular gate, rather than the occluded conformation seen in LeuT bound to its substrates^{11,48,49}. Questions extending from the substrate-bound structures involve the conformational changes required for DAT to transition to other states of the transport cycle and the roles of subsites A and B in neurotransmitter transport by mammalian DATs. Furthermore, the inhibitor-bound structures of dDAT provide a scaffold for addressing the mechanistic distinction between addictive and non-reinforcing analogues of cocaine^{25,45,50}.

Online Content Methods, along with any additional Extended Data display items, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 November 2014; accepted 23 March 2015.

Published online 11 May 2015.

- Carlsson, A., Lindqvist, M. & Magnusson, T. 3,4-Dihydroxyphenylalanine and 5-hydroxytryptophan as reserpine antagonists. *Nature* **180**, 1200 (1957).
- Twarog, B. M. & Page, I. H. Serotonin content of some mammalian tissues and urine and a method for its determination. *Am. J. Physiol.* **175**, 157–161 (1953).
- von Euler, U. Sympathin in adrenergic nerve fibres. *J. Physiol. (Lond.)* **105**, 26 (1946).
- Hertting, G. & Axelrod, J. Fate of tritiated noradrenaline at the sympathetic nerve endings. *Nature* **192**, 172–173 (1961).
- Iversen, L. L. Role of transmitter uptake mechanisms in synaptic neurotransmission. *Br. J. Pharmacol.* **41**, 571–591 (1971).
- Iversen, L. L. & Kravitz, E. A. Sodium dependence of transmitter uptake at adrenergic nerve terminals. *Mol. Pharmacol.* **2**, 360–362 (1966).
- Kristensen, A. S. et al. SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacol. Rev.* **63**, 585–640 (2011).
- Pramod, A. B., Foster, J., Carvelli, L. & Henry, L. K. SLC6 transporters: structure, function, regulation, disease association and therapeutics. *Mol. Aspects Med.* **34**, 197–219 (2013).
- Penmatsa, A., Wang, K. H. & Gouaux, E. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* **503**, 85–90 (2013).
- Koldsø, H., Christiansen, A. B., Sinning, S. & Schiøtt, B. Comparative modeling of the human monoamine transporters: similarities in substrate binding. *ACS Chem Neurosci* **4**, 295–309 (2013).

11. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* **437**, 215–223 (2005).
12. Beuming, T., Shi, L., Javitch, J. A. & Weinstein, H. A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na⁺ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function. *Mol. Pharmacol.* **70**, 1630–1642 (2006).
13. Gu, H., Wall, S. C. & Rudnick, G. Stable expression of biogenic amine transporters reveals differences in inhibitor sensitivity, kinetics, and ion dependence. *J. Biol. Chem.* **269**, 7124–7130 (1994).
14. Kilty, J. E., Lorang, D. & Amara, S. G. Cloning and expression of a cocaine-sensitive rat dopamine transporter. *Science* **254**, 578–579 (1991).
15. Kurian, M. A. *et al.* Homozygous loss-of-function mutations in the gene encoding the dopamine transporter are associated with infantile parkinsonism-dystonia. *J. Clin. Invest.* **119**, 1595–1603 (2009).
16. Giros, B., Jaber, M., Jones, S. R., Wightman, R. M. & Caron, M. G. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* **379**, 606–612 (1996).
17. Mergy, M. A. *et al.* The rare DAT coding variant Val559 perturbs DA neuron function, changes behavior, and alters *in vivo* responses to psychostimulants. *Proc. Natl Acad. Sci. USA* **111**, E4779–E4788 (2014).
18. Zaczek, R., Culp, S., Goldberg, H., McCann, D. J. & De Souza, E. B. Interactions of [3H]amphetamine with rat brain synaptosomes. I. Saturable sequestration. *J. Pharmacol. Exp. Ther.* **257**, 820–829 (1991).
19. Bönsch, H. The transport of (+)-amphetamine by the neuronal noradrenaline carrier. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **327**, 267–272 (1984).
20. Eshleman, A. J., Henningsen, R. A., Neve, K. A. & Janowsky, A. Release of dopamine via the human transporter. *Mol. Pharmacol.* **45**, 312–316 (1994).
21. Goodwin, J. S. *et al.* Amphetamine and methamphetamine differentially affect dopamine transporters *in vitro* and *in vivo*. *J. Biol. Chem.* **284**, 2978–2989 (2009).
22. Sitte, H. H. *et al.* Carrier-mediated release, transport rates, and charge transfer induced by amphetamine, tyramine, and dopamine in mammalian cells transfected with the human dopamine transporter. *J. Neurochem.* **71**, 1289–1297 (1998).
23. Sonders, M. S., Zhu, S. J., Zahniser, N. R., Kavanaugh, M. P. & Amara, S. G. Multiple ionic conductances of the human dopamine transporter: the actions of dopamine and psychostimulants. *J. Neurosci.* **17**, 960–974 (1997).
24. Wall, S. C., Gu, H. & Rudnick, G. Biogenic amine flux mediated by cloned transporters stably expressed in cultured cell lines: amphetamine specificity for inhibition and efflux. *Mol. Pharmacol.* **47**, 544–550 (1995).
25. Beuming, T. *et al.* The binding sites for cocaine and dopamine in the dopamine transporter overlap. *Nature Neurosci.* **11**, 780–789 (2008).
26. Bisgaard, H. *et al.* The binding sites for benzotropines and dopamine in the dopamine transporter overlap. *Neuropharmacology* **60**, 182–190 (2011).
27. Sørensen, G. *et al.* Neuropeptide Y Y5 receptor antagonism attenuates cocaine-induced effects in mice. *Psychopharmacology (Berl.)* **222**, 565–577 (2012).
28. Richelson, E. & Pfenning, M. Blockade by antidepressants and related compounds of biogenic amine uptake into rat brain synaptosomes: most antidepressants selectively block norepinephrine uptake. *Eur. J. Pharmacol.* **104**, 277–286 (1984).
29. Sørensen, L. *et al.* Interaction of antidepressants with the serotonin and norepinephrine transporters: mutational studies of the S1 substrate binding pocket. *J. Biol. Chem.* **287**, 43694–43707 (2012).
30. Dahal, R. A. *et al.* Computational and biochemical docking of the irreversible cocaine analog RTI 82 directly demonstrates ligand positioning in the dopamine transporter central substrate binding site. *J. Biol. Chem.* **289**, 29712–29727 (2014).
31. Pörzgen, P., Park, S. K., Hirsh, J., Sonders, M. S. & Amara, S. G. The antidepressant-sensitive dopamine transporter in *Drosophila melanogaster*: a primordial carrier for catecholamines. *Mol. Pharmacol.* **59**, 83–95 (2001).
32. Abdul-Hussein, S., Andreoli, J. & Tate, C. G. Thermostabilisation of the serotonin transporter in a cocaine-bound conformation. *J. Mol. Biol.* **425**, 2198–2207 (2013).
33. Wang, H. *et al.* Structural basis for action by diverse antidepressants on biogenic amine transporters. *Nature* **503**, 141–145 (2013).
34. Huang, X. & Zhan, C. G. How dopamine transporter interacts with dopamine: insights from molecular modeling and simulation. *Biophys. J.* **93**, 3627–3639 (2007).
35. Kitayama, S. *et al.* Dopamine transporter site-directed mutations differentially alter substrate transport and cocaine binding. *Proc. Natl Acad. Sci. USA* **89**, 7782–7785 (1992).
36. Han, D. D. & Gu, H. H. Comparison of the monoamine transporters from human and mouse in their sensitivities to psychostimulant drugs. *BMC Pharmacol.* **6**, 6 (2006).
37. Newman, A. H., Zou, M. F., Ferrer, J. V. & Javitch, J. A. [3H]MFZ 2–12: a novel radioligand for the dopamine transporter. *Bioorg. Med. Chem. Lett.* **11**, 1659–1661 (2001).
38. Kazmier, K. *et al.* Conformational dynamics of ligand-dependent alternating access in LeuT. *Nature Struct. Mol. Biol.* **21**, 472–479 (2014).
39. Krishnamurthy, H. & Gouaux, E. X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481**, 469–474 (2012).
40. Hong, W. C. & Amara, S. G. Membrane cholesterol modulates the outward facing conformation of the dopamine transporter and alters cocaine binding. *J. Biol. Chem.* **285**, 32616–32626 (2010).
41. Zhao, Y. *et al.* Substrate-modulated gating dynamics in a Na⁺-coupled neurotransmitter transporter homologue. *Nature* **474**, 109–113 (2011).
42. Malinauskaitė, L. *et al.* A mechanism for intracellular release of Na⁺ by neurotransmitter/sodium symporters. *Nature Struct. Mol. Biol.* **21**, 1006–1012 (2014).
43. Perez, C., Koshy, C., Yildiz, O. & Ziegler, C. Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* **490**, 126–130 (2012).
44. Weyand, S. *et al.* Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**, 709–713 (2008).
45. Loland, C. J. *et al.* Relationship between conformational changes in the dopamine transporter and cocaine-like subjective effects of uptake inhibitors. *Mol. Pharmacol.* **73**, 813–823 (2008).
46. Lee, S. H. *et al.* Importance of valine at position 152 for the substrate transport and 2β-carbomethoxy-3β-(4-fluorophenyl)tropane binding of dopamine transporter. *Mol. Pharmacol.* **57**, 883–889 (2000).
47. Chen, J. G., Sachpatzidis, A. & Rudnick, G. The third transmembrane domain of the serotonin transporter contains residues associated with substrate and cocaine binding. *J. Biol. Chem.* **272**, 28321–28327 (1997).
48. Singh, S. K., Piscitelli, C. L., Yamashita, A. & Gouaux, E. A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322**, 1655–1661 (2008).
49. Claxton, D. P. *et al.* Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters. *Nature Struct. Mol. Biol.* **17**, 822–829 (2010).
50. Schmitt, K. C., Rothman, R. B. & Reith, M. E. Nonclassical pharmacology of the dopamine transporter: atypical inhibitors, allosteric modulators, and partial substrates. *J. Pharmacol. Exp. Ther.* **346**, 2–10 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Coleman, C.-H. Lee, S. Mansoor and other Gouaux laboratory members for helpful discussions, L. Vaskalis for assistance with figures and H. Owen for help with manuscript preparation. We acknowledge the staff of the Berkeley Center for Structural Biology at the Advanced Light Source and the Northeastern Collaborative Access Team at the Advanced Photon Source for assistance with data collection. This work was supported by an NIMH Ruth Kirschstein postdoctoral fellowship and Brain and Behavior Research Foundation Young Investigator research award (K.H.W.), a postdoctoral fellowship from the American Heart Association (A.P.) and by the NIH (E.G.) and the Methamphetamine Abuse Research Center of OHSU (P50DA018165 to E.G.). E.G. is an investigator with the Howard Hughes Medical Institute.

Author Contributions A.P., K.H.W. and E.G. designed the project. A.P. and K.H.W. performed protein purification, crystallography and biochemical assays. A.P., K.H.W. and E.G. wrote the manuscript.

Author Information The coordinates for the structure have been deposited in the Protein Data Bank under the accession codes 4XP1, 4XP9, 4XP6, 4XPA, 4XP4, 4XP5, 4XPB, 4XPF, 4XPG, 4XPH, 4XPT (see Supplementary Table 1 for details). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.G. (gouauxe@ohsu.edu).

METHODS

Constructs. dDAT constructs employed in this study include:

subsite B mutations (*sub^B*). Denotes the addition of mutations D121G (TM helix 3) and S426M (TM helix 8) in the binding pocket.

dDAT_{wt}. Full-length dDAT without thermostabilizing mutations, fused to a C-terminal GFP tag.

dDAT_{del}. Contains an N-terminal deletion from 1–20 (Δ 1–20) and a deletion in extracellular loop 2 (EL2) from 164–191, with thermostabilizing mutations V74A and L415A.

ts² dDAT_{cryst}. Contains thermostabilizing mutations (V74A, L415A), Δ 1–20, a deletion in EL2 from Δ 164–206, and a thrombin site (LVPRGS) replacing residues 602–607 (Supplementary Table 1). Structure reported of this construct with *sub^B* mutations in complex with cocaine and RTI-55.

ts⁵ dDAT_{cryst}. Identical to *ts² dDAT_{cryst}* except that it contains three additional thermostabilizing mutations (V275A, V311A, G538L) (Supplementary Table 1). Equivalent construct containing *sub^B* mutations reported for β -CFT.

dDAT_{mfc}. Contains thermostabilizing mutations (V74A, L415A), Δ 1–20, a modified deletion in EL2, that is, Δ 162–202, and a thrombin site replacing residues 602–607 (Supplementary Table 1). Structures in complex with cocaine, RTI-55, D-amphetamine, (+)-methamphetamine, DA and 3,4-dichlorophenethylamine (DCP) are reported. Structure of dDAT_{mfc} containing *sub^B* mutations reported in complex with DCP.

dDAT_{mfc} 201. Identical to dDAT_{mfc} except the deletion in EL2 is from Δ 162–201. (Supplementary Table 1). Structure of dDAT_{mfc} 201 containing *sub^B* mutations reported in complex with DCP.

Expression and purification. The dDAT constructs were expressed as C-terminal green fluorescent protein (GFP)–His₈ fusions using baculovirus-mediated transduction of mammalian HEK-293S GnTI[−] cells^{51,52}. Membranes harvested from cells post-infection were homogenized with 1× TBS (20 mM Tris pH 8.0, 100 mM NaCl) and solubilized with a final concentration of 20 mM *n*-dodecyl β -D-maltoside (DDM) and 4 mM cholesteryl hemisuccinate (CHS) in 1× TBS. Detergent-solubilized material was incubated with cobalt-charged metal affinity resin and eluted with 1× TBS containing 1 mM DDM and 0.2 mM CHS along with 80 mM imidazole (pH 8.0). The GFP–His₈ tag was removed using thrombin digestion followed by concentrating the metal ion affinity purified protein. The thrombin-digested protein was subjected to size exclusion chromatography through a Superdex 200 10/300 column pre-equilibrated with buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 4 mM decyl β -D-maltoside, 0.2 mM CHS and 0.001% (w/v) 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE). Peak fractions greater than 0.5 mg ml^{−1} were collected and pooled together. Ascorbic acid (25 mM) was added to the protein solution used to crystallize the DA–dDAT_{mfc} complex, to serve as antioxidant. All procedures were carried out at 4 °C.

Fab complexation and crystallization. Antibody fragment (Fab) 9D5 was used to complex with the protein at a molar ratio of 1.2 (Fab):1 (protein). The Fab–DAT complex solution was incubated with 1–2 mg solid drug for 30 min on ice followed by concentration in a 100 kDa cutoff concentrator to 3.5–5 mg ml^{−1}. The concentrated protein was spun down to remove excess drug and insoluble aggregates and plates were set up by hanging drop vapour diffusion. Crystals of Fab–DAT complex grew primarily in conditions containing PEG 400 or PEG 350 monomethyl ether or PEG 600 as the precipitant (Extended Data Table 2b). The pH range of crystallization was 8.0–9.0 for dDAT_{cryst}-based constructs and in the range of 6.5–8.0 with dDAT_{mfc}-based constructs (Extended Data Table 2b). Crystals of dDAT_{mfc} were primarily obtained by streak seeding with a cat whisker dipped in crystals formed with *sub^B* containing constructs, 2–5 days after drops were set up. All crystals were grown at 4 °C.

Data collection and structure refinement. Crystals were directly flash-cooled in liquid nitrogen when the PEG 400 concentration in the mother liquor exceeded 36%. For crystals grown in wells containing less than 36% PEG 400, crystals were transferred into cryoprotection solution identical to the reservoir solution but with 40% PEG 400. In conditions with 30–34% PEG 600 as the primary precipitant, 10% of ethylene glycol was added to provide additional cryoprotection. Data were collected either at ALS (5.0.2; 8.2.1) or APS (24-IDC and IDE). Anomalous data for iodine containing RTI-55 complexed with dDAT was collected at 1.6 Å as described in the crystallographic data table. Data were processed using either HKL2000^{52,53} or XDS⁵⁴. Molecular replacement was carried out for all data sets using coordinates 4M48 with Fab 9D5 and dDAT_{cryst} used as independent search models, using PHASER in the PHENIX software suite^{55,56}. Iterative cycles of refinement and manual model building were carried out using PHENIX and COOT⁵⁷, respectively, until the models converged to acceptable levels of R-factors and stereochemistry.

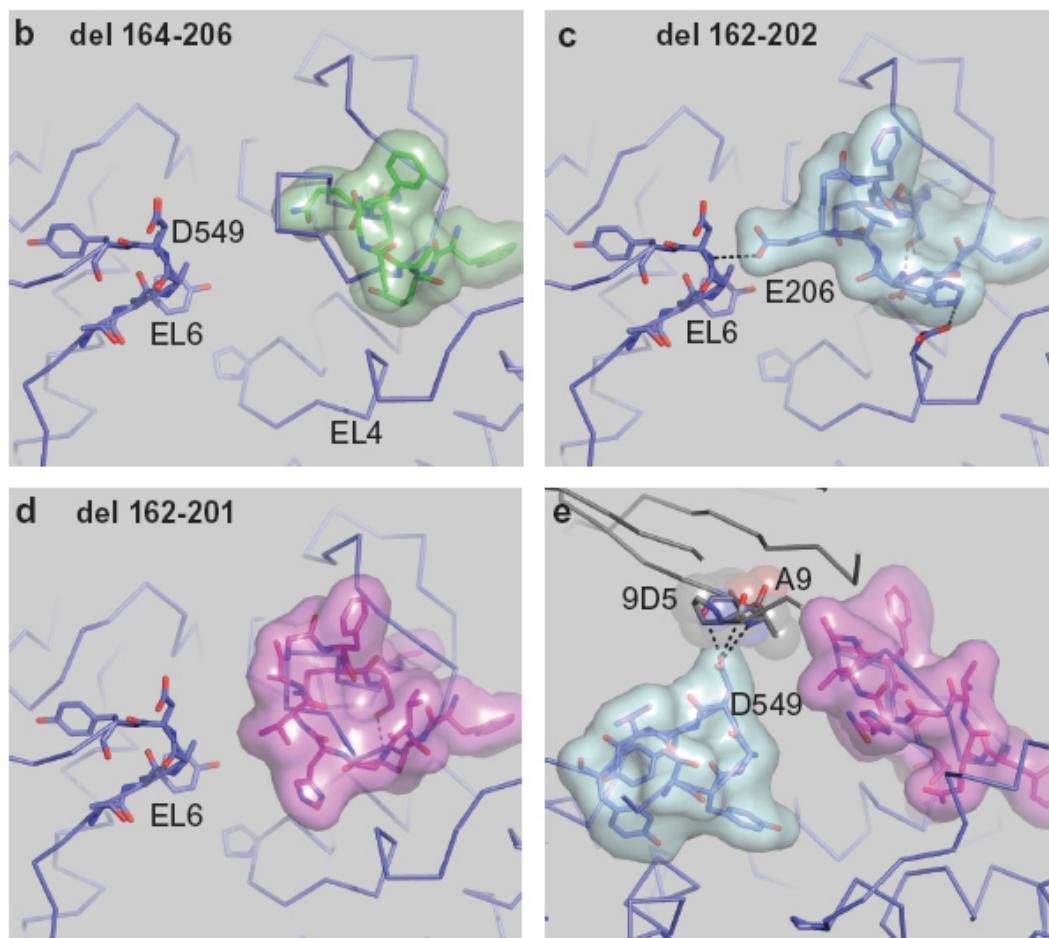
Radiolabel binding and uptake assays. All binding assays were carried out by scintillation proximity assay (SPA) method⁵⁸. Reactions contained 5–20 nM protein, 0.5 mg ml^{−1} Cu-YSi beads, SEC buffer, and [³H]nisoxetine from 0.1 to 300 nM for saturation binding assays. Competition binding assays were done with 30 nM [³H]nisoxetine and increasing concentrations of unlabelled competitor. *K_i* values were estimated from IC₅₀ values using the Cheng–Prusoff equation. Fits were plotted using Graphpad Prism v4.0.

For uptake assays HEK 293S cells were infected with baculovirus expressing dDAT_{mfc} or dDAT_{wt}, and sodium butyrate was added to 10 mM 8–12 h post-infection. At 24 h, infected cells were adhered to Cytostar-T plates for 2 h, after which the culture media were replaced with uptake buffer (20 mM HEPES pH 7.4, 120 mM NaCl, 5 mM KCl, 2.5 mM CaCl₂, 1.2 mM MgSO₄, 10 mM D-glucose, 1 mM tropolone, 1 mM L-ascorbic acid^{59,60}). Uptake assays were carried out using [¹⁴C]DA over a range of 0.3–60 μ M in 100 μ l total volume. Samples were read every five minutes for a time course over twenty-five minutes. The linear initial rate of uptake was plotted in the presence and absence of 10 μ M nortriptyline to calculate the specific uptake rate. Data were fitted to a standard Michaelis–Menten equation to obtain *K_M* and *V_{max}* values. The significance of specific uptake was assessed at each concentration of DA using a two-tailed Welch's *t*-test with 2 degrees of freedom.

- Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99**, 13419–13424 (2002).
- Goehring, A. *et al.* Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nature Protocols* **9**, 2574–2585 (2014).
- Otwinowski, Z., & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Quick, M. & Javitch, J. A. Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc. Natl Acad. Sci. USA* **104**, 3603–3608 (2007).
- Eshleman, A. J. *et al.* Metabolism of catecholamines by catechol-O-methyltransferase in cells expressing recombinant catecholamine transporters. *J. Neurochem.* **69**, 1459–1466 (1997).
- Dehnes, Y. *et al.* Conformational changes in dopamine transporter intracellular regions upon cocaine binding and dopamine translocation. *Neurochem. Int.* **73**, 4–15 (2014).

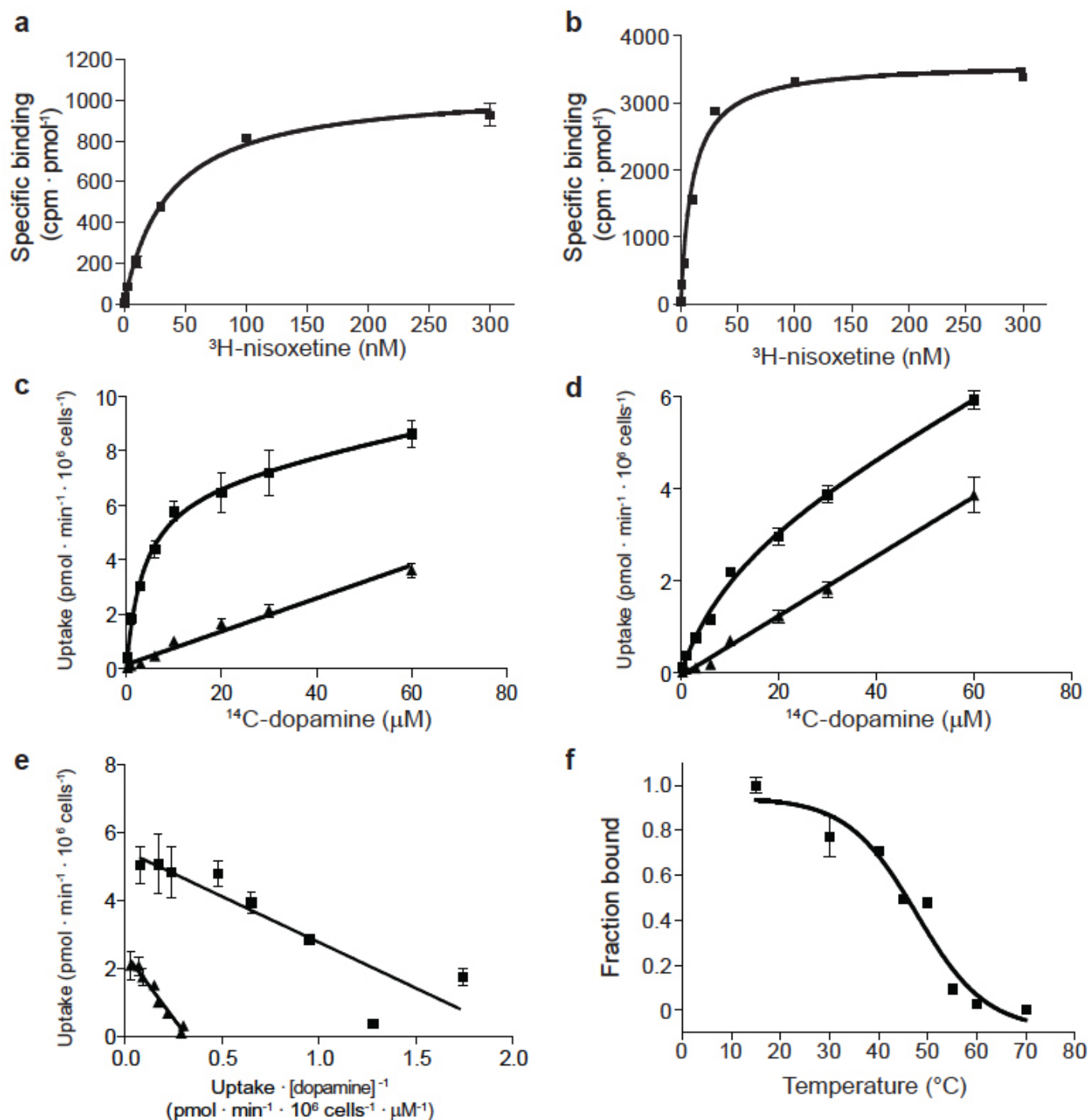
a**Extracellular Loop (EL) 2 deletions**

wt dDAT	WNTPNCRPFESQNASRVVPVIGNYSDLYAMGNQSLLYNETYMNGSSLDTSVAGHVEGFQSA
del164-206 (DAT _{cryst})	WNTPNCRPFESQ-----163-----207-----GFQSA
del162-202 (DAT _{mfc})	WNTPNCRPFE-----161-----203-----GHVEGFQSA
del162-201	WNTPNCRPFE-----161-----202-----VGHVEGFQSA

**Extended Data Figure 1 | Design of the minimal functional construct.**

a, Thermostabilizing (ts) mutations V275A, V311A, G538L were removed. Modification of the EL2 deletion from 164–206 to 162–202, which recovered transport activity. The del 162–201 construct has robust dopamine uptake activity. **b**, Structural organization of EL2 regions. Organization of dDAT_{cryst} with a deletion of region 164–206 depicted as green surface. **c**, EL2 structure in

dDAT_{mfc} with the deletion 162–202 depicted as cyan surface showing contacts between EL2 and EL6. **d**, EL2 organization in the construct with a deletion from 162–201 depicted as magenta surface. **e**, Fab 9D5 interferes with the interaction between EL2 and EL6 in the crystal lattice, with loops depicted as magenta and cyan surfaces, respectively. Fab disrupts the EL organization in all structures. The del 162–201 sub^B structure is shown.



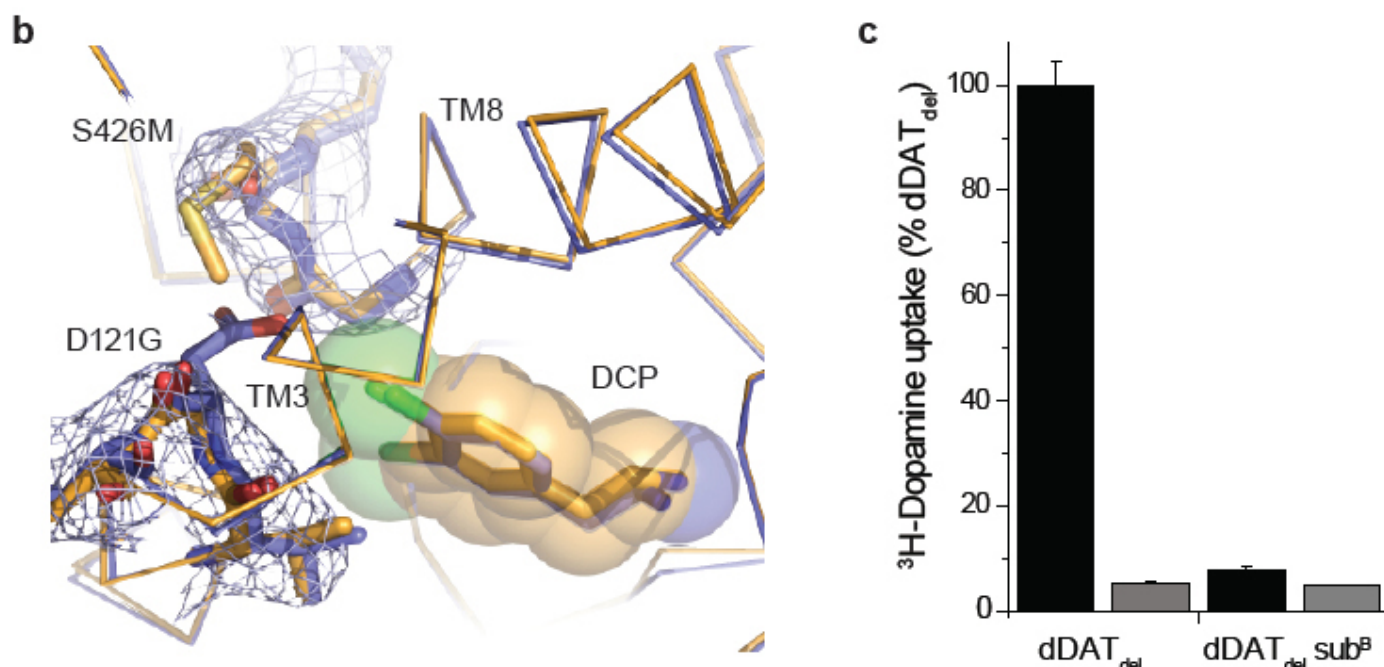
Extended Data Figure 2 | Measurement of dissociation constants using purified dDATmfc protein, and dopamine uptake in whole cells.

a, dDAT_{mfc} binds [³H]nisoxetine with a K_d of 36 ± 3 nM (s.e.m.). **b**, dDAT_{mfc} with sub^B mutations binds [³H]nisoxetine with a K_d of 10 ± 1 nM (s.e.m.). **c**, **d**, Michaelis-Menten plots of [¹⁴C]dopamine uptake by HEK293S cells expressing dDAT_{wt} or dDAT_{mfc}, respectively, which yielded a K_M of 2.1 ± 0.7 μM and V_{max} of 4.5 ± 0.4 pmol min⁻¹ per 10⁶ cells for dDAT_{wt} and a K_M of 8.2 ± 2.3 μM and V_{max} of 2.4 ± 0.2 pmol min⁻¹ per 10⁶ cells for dDAT_{mfc} (s.e.m.). One representative plot of total and background counts (in the presence of 10 μM nortriptyline) is shown of two experimental trials as squares and triangles, respectively. Data points and error bars show the average and standard deviation, respectively, of technical replicates ($n = 3$). Welch's

t -test indicates that the specific uptake signal at each concentration of dopamine is significant with a two-tailed P value < 0.02 . **e**, Eadie-Hofstee plot of specific dopamine uptake shown in Fig. 1a and panels **c** and **d** of this figure. Data for dDAT_{wt} and dDAT_{mfc} are shown as squares and triangles, respectively, and error bars denote s.d. of technical replicates ($n = 3$). **f**, The thermal melting curve of dDAT_{mfc} solubilized from HEK293S membranes in the presence of 100 nM [³H]nisoxetine exhibits a melting temperature of 48 ± 2 °C (s.e.m.). The fraction bound describes the signal remaining after incubation at the specified temperature for 10 min, normalized to the signal at 4 °C. Data points show the mean values for one experimental trial, and error bars show the s.d. of technical replicates ($n = 3$).

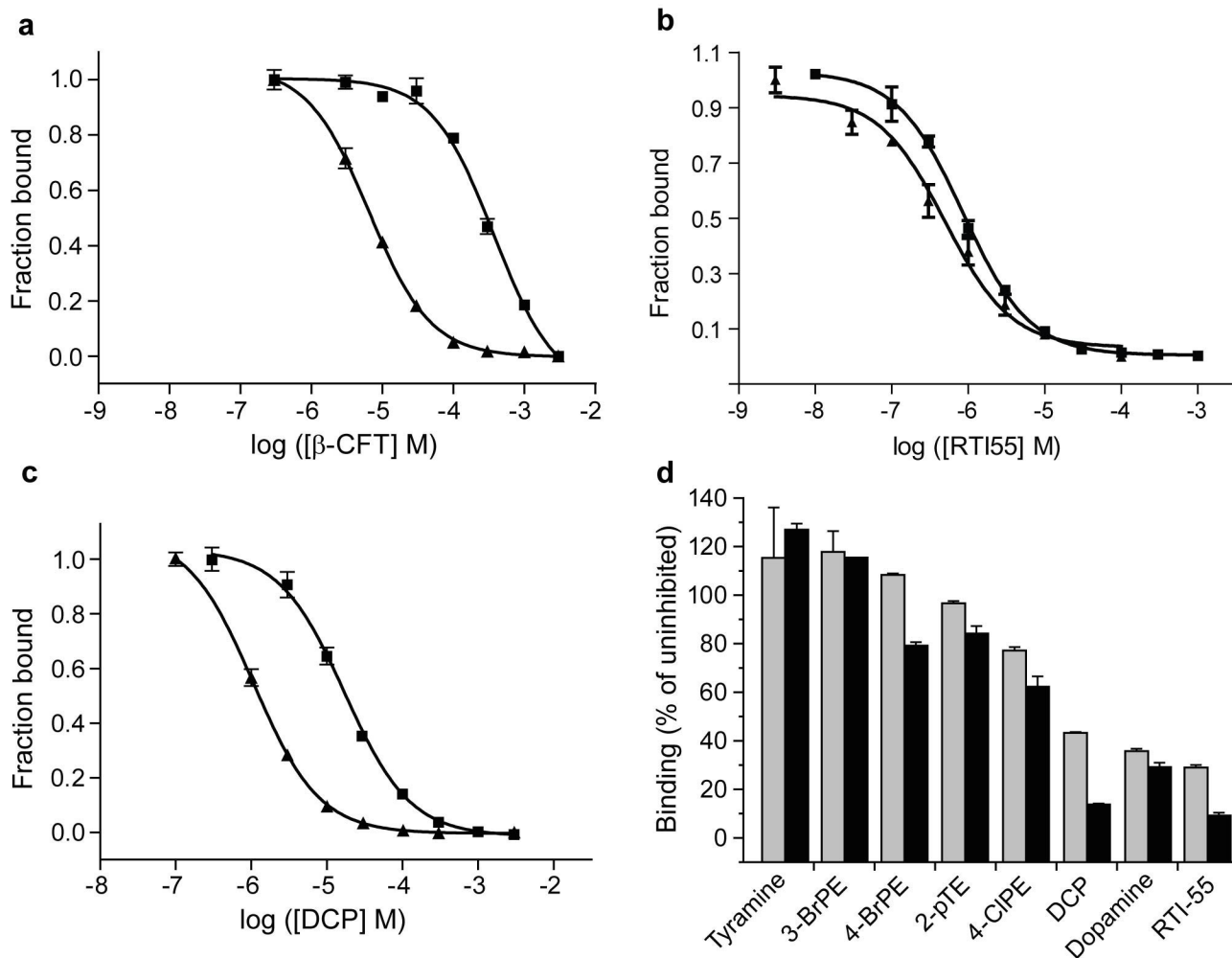
a **Locations of Subsite B mutations**

	D121G (TM3)	S426M (TM8)
dDAT	116-IAFYVDFYINV-126	421-SSFGGSEAIITALSD-435
hDAT	148-ISLYVGFFYNV-158	422-SAMGGMESVITGLID-436
hNET	144-IALYVGFYINV-154	419-SSMGGMEAVITGLAD-433
hSERT	168-IAFYIASYYNT-178	438-STFAGLEGVITAVLD-452



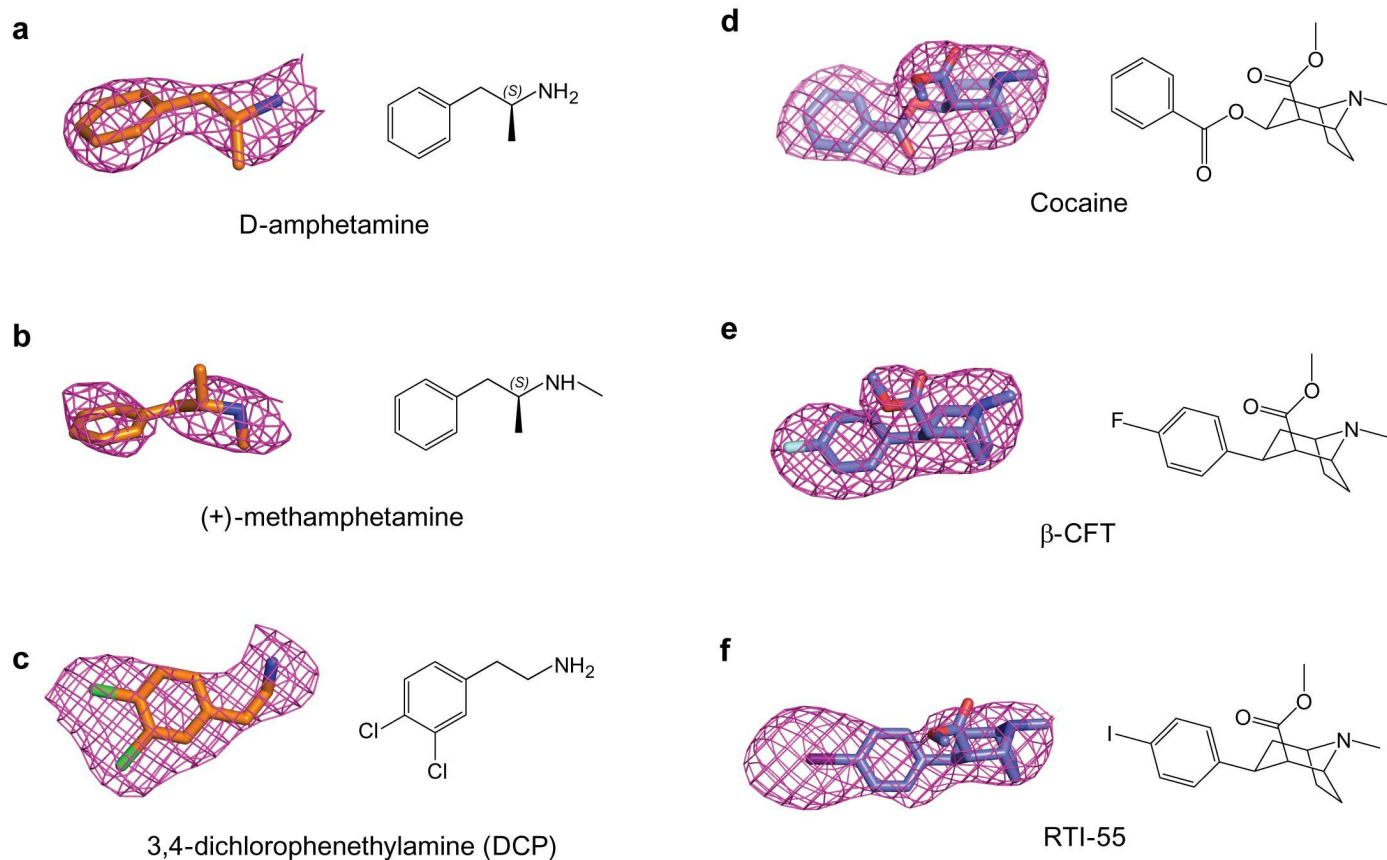
Extended Data Figure 3 | Mutagenesis and effects of DAT subsite B.
a, Sequence alignment of subsite B regions for dDAT and human NSS orthologues. **b**, $2F_o - F_c$ density contoured at 0.9σ around the vicinity of the D121G (TM3) and S426M mutations (TM8). **c**, Abrogation of dopamine

transport activity by dDAT_{wt} bearing both subsite B mutations in infected HEK293S cells. Data show the average uptake and error bars show the data range of technical duplicates for a single trial. Reactions were performed without and with 100 μ M desipramine in black and grey bars, respectively.

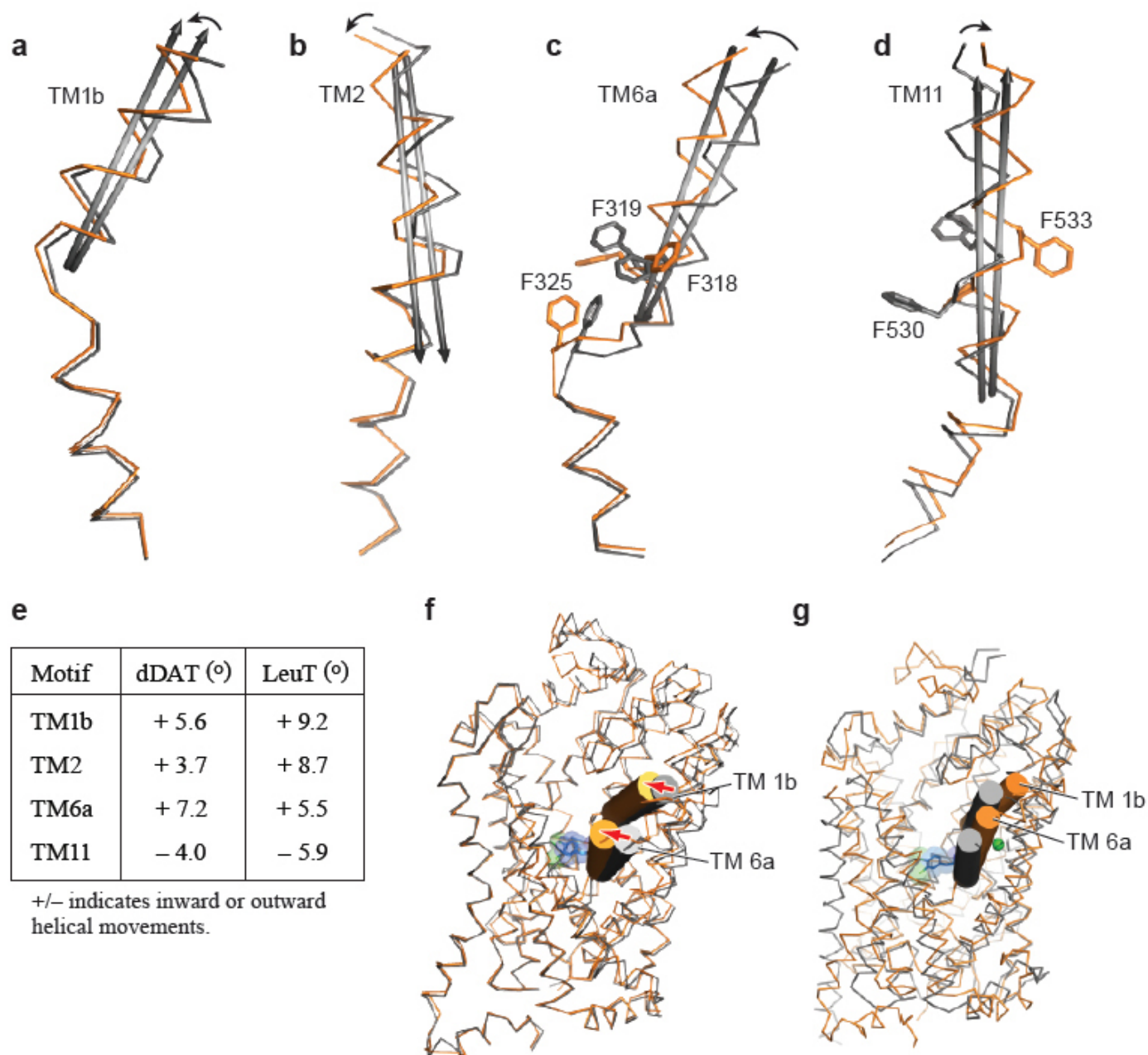


Extended Data Figure 4 | Measurement of inhibition constants using purified dDAT protein. **a–c**, Inhibition of [^3H]nisoxetine binding to dDAT_{mfc} (squares) and dDAT_{mfc} sub^B (triangles). K_i inhibition constants for dDAT_{mfc} and dDAT_{mfc} sub^B are, respectively, $98 \pm 4 \mu\text{M}$, and $1.4 \pm 0.1 \mu\text{M}$, (**a**, β -CFT), $371 \pm 25 \text{ nM}$ and $271 \pm 59 \text{ nM}$ (**b**, RTI-55), and $4.5 \pm 0.3 \mu\text{M}$, and $267 \pm 20 \text{ nM}$ (**c**, DCP). All errors are s.e.m. One representative trial of two is shown for all experiments in panels **a–c**, and data points and error bars denote

the average values for fraction bound and standard deviation, respectively, for technical replicates ($n = 3$). **d**, Inhibition of [^3H]nisoxetine (50 nM) binding to dDAT_{del} by 1 and 10 μM unlabelled compound (grey and black bars, respectively). Error bars show the data range of technical replicates ($n = 2$). Abbreviations: 3-BrPE, 3-bromophenethylamine; 4-BrPE, 4-bromophenethylamine; 2-pTE, 2-(pTolyl)ethylamine; 4-ClPE, 4-chlorophenethylamine; DCP, 3,4-dichlorophenethylamine.

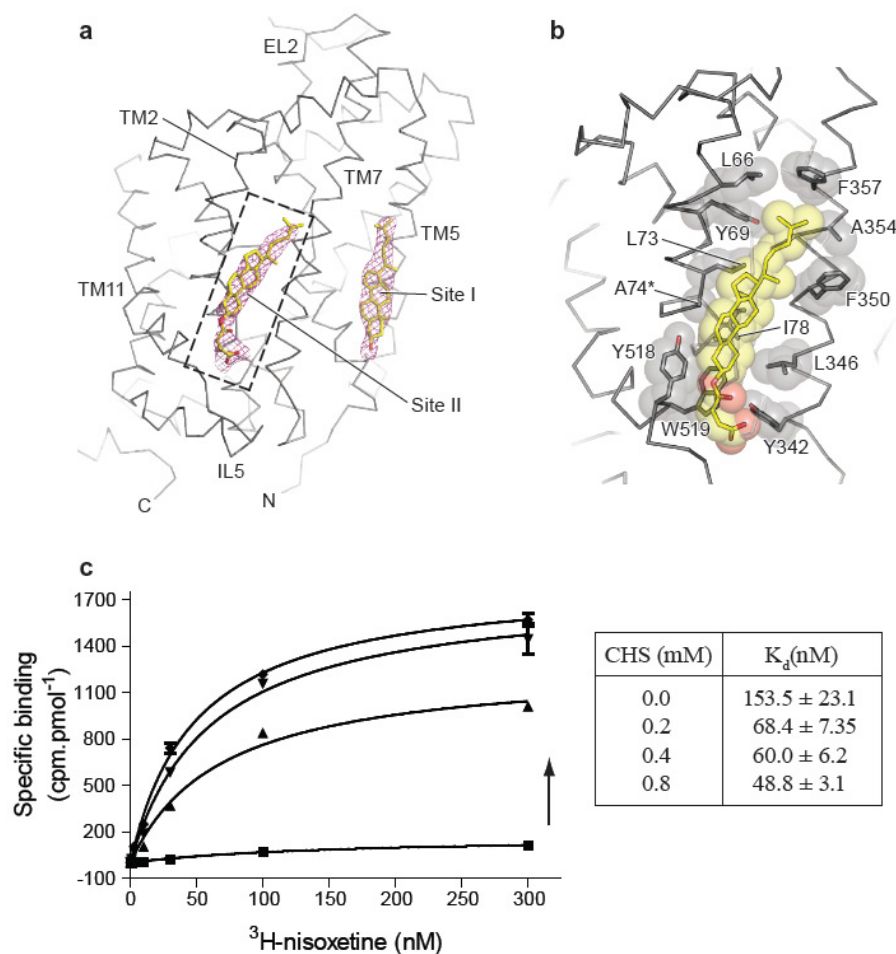


Extended Data Figure 5 | $F_o - F_c$ densities for ligands complexed with dDAT. **a**, D-amphetamine (2.4σ); **b**, (+)-methamphetamine (1.8σ); **c**, DCP (2.2σ); **d**, cocaine (2.2σ); **e**, β -CFT (2.2σ); **f**, RTI-55 (2.6σ).



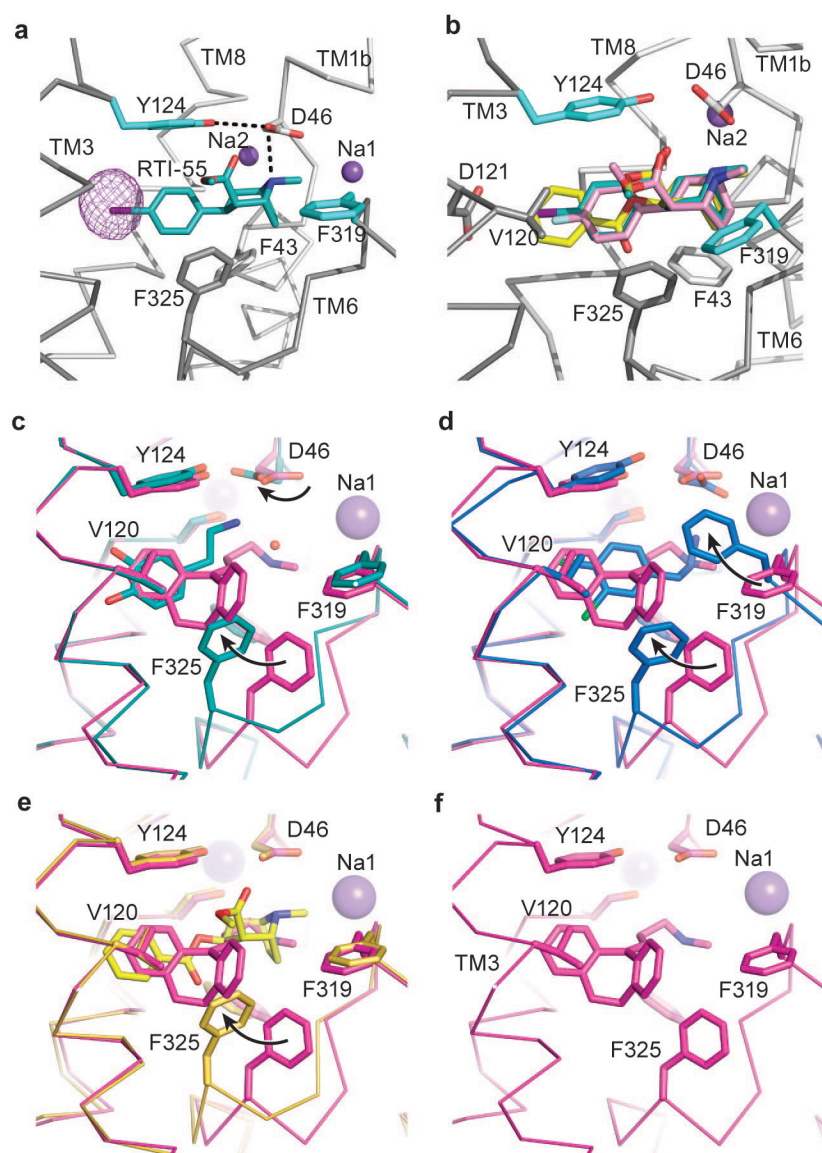
Extended Data Figure 6 | Helical movements in dDAT_{mfc} upon binding to substrate analogue DCP (orange) and inhibitor nortriptyline (grey). **a–d**, Helices undergoing maximal shifts are **a**, TM1b; **b**, TM2; **c**, TM6a; **d**, TM11. Arrows in black represent direction of shift. **e**, Table comparing angular shifts between nortriptyline–dDAT_{cryst} (PDB ID 4M48) and DCP–dDAT_{mfc} structures in column one, and between the outward-open Trp–LeuT (PDB ID 3F3A) and outward-occluded Leu–LeuT (PDB ID 2A65) structures in

column two. **f**, Superposition of the outward open state of nortriptyline–dDAT_{cryst} (PDB ID 4M48) and DCP–dDAT_{mfc} structures in grey and orange ribbon, respectively. Extracellular gating TMs 1b and 6a are shown as cylinders. Arrows in red indicate inward movement of TMs 1b and 6a. **g**, Superposition of the occluded state of LeuT (PDB ID 2A65) and DCP–dDAT_{mfc} structures in grey and orange ribbon, respectively.



Extended Data Figure 7 | Cholesterol binding sites in DAT. **a**, Cholesterol binding sites seen on the dDAT surface corresponding to the inner leaflet of the plasma membrane, with a second novel cholesterol site into which a cholesteryl hemisuccinate (CHS) could be modelled. $F_o - F_c$ densities for cholesterol contoured at 2.0σ . **b**, Close-up view of cholesterol site II at the junction of TM2, TM7 and TM11 interacting with multiple hydrophobic residues. Asterisk

denotes thermostabilizing mutant V74A. **c**, Effect of CHS concentration on [³H]nisoxetine binding to DAT_{mfc} construct. Graph depicts one representative trial of two independent experiments, and total and background counts were measured using technical replicates ($n = 3$) for each binding curve at each CHS concentration. Arrow represents increasing concentration of CHS. Error bars represent s.d.



Extended Data Figure 8 | Analogues of cocaine and binding site comparisons. **a**, The position of RTI-55 in the binding pocket with anomalous difference density for iodide displayed as purple mesh and contoured at 4σ . **b**, Superposition of cocaine, β -CFT, and RTI-55 using the RTI-55-dDAT_{mfc} structure. Ligands are shown as sticks and coloured yellow (cocaine), pink (β -CFT), and teal (RTI-55). Sodium ions are shown as purple spheres.

c–f, Residues that line the binding pocket are superposed between the nortriptyline-dDAT_{cryst} (magenta, PDB ID 4M48) and those of **c**, DA-dDAT_{mfc} (cyan), **d**, DCP-dDAT_{mfc} (marine), **e**, cocaine-dDAT_{mfc} (yellow). **f**, Organization of S1 binding site in complex with nortriptyline (PDB ID 4M48). Black arrows describe the change in rotamers and positions of D46, F319, and F325 compared to the nortriptyline-bound structure.

Extended Data Table 1 | Superposition statistics of dDAT structures

	ntt ts ⁵ dDAT _{cryst}	cocaine dDAT _{mfc}	β-CFT ts ⁵ dDAT _{cryst} sub ^B	RTI-55 dDAT _{mfc} sub ^B	RTI-55 dDAT _{mfc}	DCP dDAT _{mfc} sub ^B	DCP dDAT _{mfc}	DA dDAT _{mfc}	meth dDAT _{mfc}	amph dDAT _{mfc}
ntt ts ⁵ dDAT _{cryst}	—	0.46	0.37	0.52	0.70	0.74	0.76	0.53	0.53	0.48
cocaine dDAT _{mfc}	—	—	0.27	0.46	0.67	0.74	0.79	0.52	0.54	0.37
β-CFT ts ⁵ dDAT _{cryst} sub ^B			—	0.43	0.64	0.70	0.73	0.50	0.50	0.52
RTI-55 ts ² dDAT _{cryst} sub ^B				—	0.73	0.75	0.82	0.55	0.59	0.60
RTI-55 dDAT _{mfc}					—	0.83	0.64	0.97	0.53	0.58
DCP dDAT _{mfc} sub ^B						—	0.41	0.75	0.57	0.61
DCP dDAT _{mfc}							—	0.64	0.64	0.63
DA dDAT _{mfc}								—	0.64	0.41
meth dDAT _{mfc}									—	0.35
amph dDAT _{mfc}										—

Extended Data Table 2 | a, Ligand surface and interface areas*; b, crystallization conditions for ligand–DAT complexes

a

Ligand	Total surface area (Å ²)	Buried Surface area (Å ²)	% Buried
Nortriptyline (4M48)	475.9	448.4	94.2
Cocaine	484.7	447	92.2
RTI-55	466.2	416.7	89.4
β-CFT	439.2	403.1	91.8
dopamine	311.9	279.1	89.5
DCP	337.3	321.5	95.3
methamphetamine	325.7	301	92.4
D-amphetamine	308.3	284.5	92.3

b

	Ligand	Construct	Condition
1	Cocaine	dDAT _{mfc}	PEG 400 37%, Na MES pH 6.8 (0.1M)
2	RTI-55	dDAT _{mfc}	PEG 400 41%, Tris pH 8.0 (0.1M)
3	Methamphetamine	dDAT _{mfc}	PEG 400 38%, Tris pH 8.0 (0.1M)
4	D-amphetamine	dDAT _{mfc}	PEG 600 36%, MOPS pH 7.0 (0.1M)
5	DA (dopamine)	dDAT _{mfc}	PEG 600 31%, Tris pH 8.0 (0.1M)
6	DCP	dDAT _{mfc}	PEG 400 38%, Tris pH 8.0 (0.1M)
7	Cocaine	dDAT _{cryst} ts ² sub ^B	PEG 400 38%, Tris-Bicine pH 8.5 (0.1M)
8	RTI-55	dDAT _{cryst} ts ² sub ^B	PEG 400 39%, Bicine pH 8.8 (0.1M)
9	β-CFT	dDAT _{cryst} sub ^B	PEG 400 38%, Bicine pH 8.8 (0.1M)
10	DCP	dDAT _{mfc} sub ^B	PEG 400 33%, Na MES pH 6.5 (0.1M)
11	DCP	dDAT _{mfc} 201 sub ^B	PEG400 34%, Na MES pH 6.5 (0.1M)

*Surface areas calculated by PDBePISA. (<http://www.ebi.ac.uk/pdbe/pisa>)

A strong ultraviolet pulse from a newborn type Ia supernova

Yi Cao¹, S. R. Kulkarni^{1,2}, D. Andrew Howell^{3,4}, Avishay Gal-Yam⁵, Mansi M. Kasliwal⁶, Stefano Valenti^{3,4}, J. Johansson⁷, R. Amanullah⁷, A. Goobar⁷, J. Sollerman⁸, F. Taddia⁸, Assaf Horesh⁵, Ilan Sagiv⁵, S. Bradley Cenko⁹, Peter E. Nugent^{10,11}, Iair Arcavi^{3,12}, Jason Surace¹³, P. R. Woźniak¹⁴, Daniela I. Moody¹⁴, Umaa D. Rebbapragada¹⁵, Brian D. Bue¹⁵ & Neil Gehrels⁹

Type Ia supernovae¹ are destructive explosions of carbon-oxygen white dwarfs^{2,3}. Although they are used empirically to measure cosmological distances^{4–6}, the nature of their progenitors remains mysterious³. One of the leading progenitor models, called the single degenerate channel, hypothesizes that a white dwarf accretes matter from a companion star and the resulting increase in its central pressure and temperature ignites thermonuclear explosion^{3,7,8}. Here we report observations with the Swift Space Telescope of strong but declining ultraviolet emission from a type Ia supernova within four days of its explosion. This emission is consistent with theoretical expectations of collision between material ejected by the supernova and a companion star⁹, and therefore provides evidence that some type Ia supernovae arise from the single degenerate channel.

On UTC (Coordinated Universal Time) 2014 May 3.29 the intermediate Palomar Transient Factory (iPTF, a wide-field survey designed to search for optical transient and variable sources)¹⁰ discovered an optical transient, internally designated as iPTF14atg, in the apparent vicinity of the galaxy IC 831 at a distance of 93.7 Mpc (see Methods subsection ‘Discovery’). No activity had been detected at the same location in the images taken on the previous night and earlier, indicating that the supernova probably exploded between May 2.29 and 3.29. Our follow-up spectroscopic campaign (see Extended Data Table 1 for the observation log) established that iPTF14atg was a type Ia supernova.

Upon discovery we triggered observations with the Ultraviolet/Optical Telescope (UVOT) and the X-ray Telescope (XRT) onboard the Swift space observatory¹¹ (observation and data reduction is detailed in Methods subsection ‘Data acquisition’; raw measurements are shown in Extended Data Table 2). As can be seen in Fig. 1, the ultraviolet brightness of iPTF14atg declined substantially in the first two observations. A rough energy flux measure in the ultraviolet band is provided by $\nu f_\nu \approx 3 \times 10^{-13}$ erg cm⁻² s⁻¹ in the ‘uvm2’ band. Starting from the third epoch, the ultraviolet and optical emission began to rise again in a manner similar to that seen in other type Ia supernovae. The XRT did not detect any X-ray signal at any epoch (Methods subsection ‘Data acquisition’). We thus conclude that iPTF14atg emitted a pulse of radiation primarily in the ultraviolet band. This pulse with an observed luminosity of $L_{UV} \approx 3 \times 10^{-41}$ erg s⁻¹ was probably already declining by the first epoch of the Swift observations (within four days of its explosion).

Figure 1 also illustrates that such an early ultraviolet pulse from a type Ia supernova within four days of its explosion is unprecedented^{12,13}. We now seek an explanation for this early ultraviolet emis-

sion. As detailed in Methods subsection ‘Spherical models for the early ultraviolet pulse’, we explored models in which the ultraviolet emission is spherically symmetric with the supernova explosion (such as shock cooling and circumstellar interaction). These models are unable to explain the observed ultraviolet pulse. Therefore we turn to asymmetric models in which the ultraviolet emission comes from particular directions.

A reasonable physical model is ultraviolet emission arising in the ejecta as the ejecta encounters a companion^{9,14}. When the rapidly moving ejecta slams into the companion, a strong reverse shock is generated in the ejecta that heats up the surrounding material. Thermal radiation from the hot material, which peaks in the ultraviolet part of the spectrum, can then be seen for a few days until the fast-moving ejecta engulfs the companion and hides the reverse shock region. We compare a semi-analytical model⁹ to the Swift/UVOT lightcurves. For simplicity, we fix the explosion date at 2014 May 3. We assume that the exploding white dwarf is close to the Chandrasekhar mass limit (1.4 solar masses) and that the supernova explosion energy is 10⁵¹ erg. These values lead to a mean expansion velocity of 10⁴ km s⁻¹ for the ejecta. Since the temperature at the collision location is so high that most atoms are ionized, the opacity is probably dominated by electron scattering. To further simplify the case, we assume that the emission from the reverse shock region is blackbody and isotropic. In order to explain the ultraviolet lightcurves, the companion star should be located 60 solar radii away from the white dwarf (model A; black dashed curves in Fig. 1).

There are several caveats in this simple semi-analytical model. First, the model parameters are degenerate. For example, if we reduce the supernova energy by a factor of two and increase the binary separation to 90 solar radii, the model lightcurve can still account for the observed ultraviolet luminosities (model B; blue dashed curves in Fig. 1). Second, the emission from the reverse shock region is not isotropic. The ultraviolet photons can only easily escape through the conical hole carved out by the companion star and therefore the emission is more concentrated in this direction. Third, the actual explosion date is not well constrained, so that when exactly the companion collision happened is not clear. Our multi-wavelength observations soon after discovery of the supernova provide a good data set for detailed modelling.

We also construct the spectral energy distribution from the photometry and spectrum of iPTF14atg obtained on the same day of the first UVOT epoch and compare it with the blackbody spectra derived from models A and B. As can be seen in Fig. 2, the model blackbody spectra are consistent with the overall shape of the spectral energy distribution, indicating that the emitting regions can be approximated

¹Astronomy Department, California Institute of Technology, Pasadena, California 91125, USA. ²Caltech Optical Observatories, California Institute of Technology, Pasadena, California 91125, USA. ³Las Cumbres Observatory Global Telescope Network, 6740 Cortona Drive, Suite 102, Goleta, California 93117, USA. ⁴Department of Physics, University of California, Santa Barbara, California 93106, USA. ⁵Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel. ⁶Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, California 91101, USA. ⁷The Oskar Klein Centre, Department of Physics, Stockholm University, SE-10691 Stockholm, Sweden. ⁸The Oskar Klein Centre, Department of Astronomy, Stockholm University, SE-10691 Stockholm, Sweden. ⁹Astrophysics Science Division, NASA Goddard Space Flight Center, Mail Code 661, Greenbelt, Maryland 20771, USA. ¹⁰Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ¹¹Astronomy Department, University of California, Berkeley, 501 Campbell Hall, Berkeley, California 94720, USA. ¹²Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA. ¹³Spitzer Science Center, California Institute of Technology, Pasadena, California 91125, USA. ¹⁴Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. ¹⁵Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA.

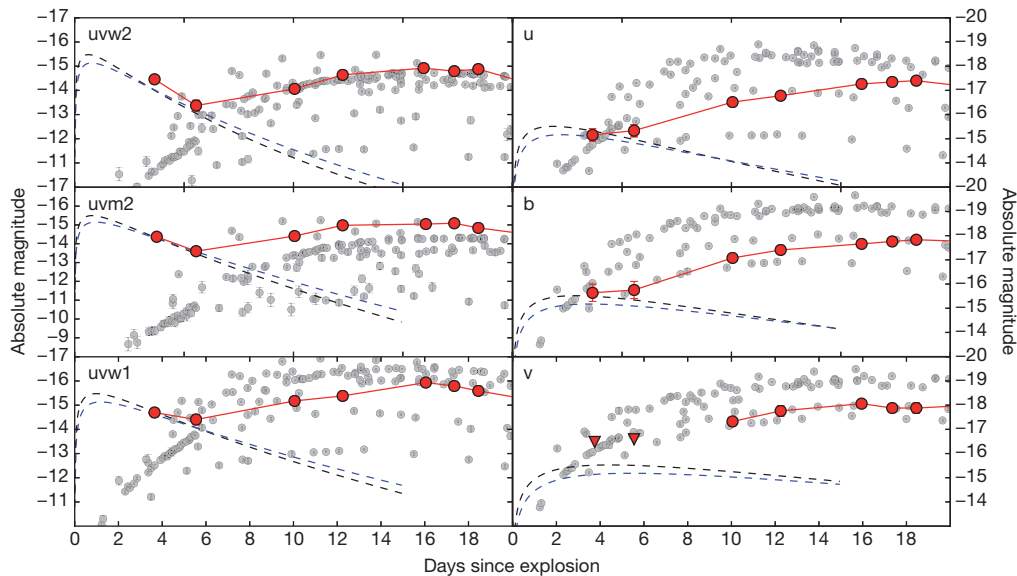


Figure 1 | Swift/UVOT lightcurves of iPTF14atg. iPTF14atg lightcurves are shown with red circles and lines and are compared with those of other type Ia supernovae (grey circles). The magnitudes are in the AB system. The 1σ error bars include both statistical and systematic uncertainties in measurements. Lightcurves of other supernovae and their explosion dates are taken from

previous studies^{13,26}. In each of the three ultraviolet bands (uvw2, uvm2 and uvw1), iPTF14atg stands out as exhibiting a decaying flux at early times. The blue and black dashed curves show two theoretical lightcurves derived from companion interaction models⁹.

by a blackbody with a temperature of 11,000 K and a radius of 6,000 solar radii.

Next, given the diversity of type Ia supernovae, we investigate the specifics of iPTF14atg using its multi-band lightcurves (Fig. 3) and spectra (Fig. 4). First, the existence of Si II and S II absorption features in the pre-maximum spectra indicates that iPTF14atg is spectroscopically a type Ia supernova¹. Second, iPTF14atg, with a peak absolute magnitude of -17.9 mag in the B band, is 1.4 magnitudes fainter than normal type Ia supernovae, which are used as cosmological distance indicators¹⁵. Subluminous type Ia supernovae belong to three major families, with prototypical events being SN 1991bg¹⁶, SN 2002cx^{17,18}, and SN 2002es¹⁹. A comparative analysis of lightcurves and spectra between iPTF14atg and the three families (detailed in Methods subsection ‘Supernova specification’) shows that iPTF14atg is more luminous than SN 1991bg and evolves more slowly than SN 1991bg

in both the rise and decline phases. The expansion velocity of iPTF14atg estimated from absorption lines is systematically lower than that of SN 1991bg. SN 2002cx and iPTF14atg have similar lightcurves, but iPTF14atg shows deep silicon absorption features in the pre-maximum spectra that are not seen in SN 2002cx and the post-maximum absorption features of iPTF14atg are generally weaker than those seen in SN 2002es. We have only limited knowledge about the evolution of SN 2002es. Despite the fact that SN 2002es is one magnitude brighter at peak than iPTF14atg and that the lightcurve of SN 2002es shows an accelerating decline about 30 days after its peak, which is not seen in iPTF14atg, iPTF14atg shows a reasonable match to SN 2002es in both lightcurve shape and spectra with higher line velocities. In addition, the host galaxy IC 831 of iPTF14atg is an early-type galaxy. This is consistent with the host galaxies of known events similar to SN 2002es, while the majority of events like SN 2002cx occur in late-type

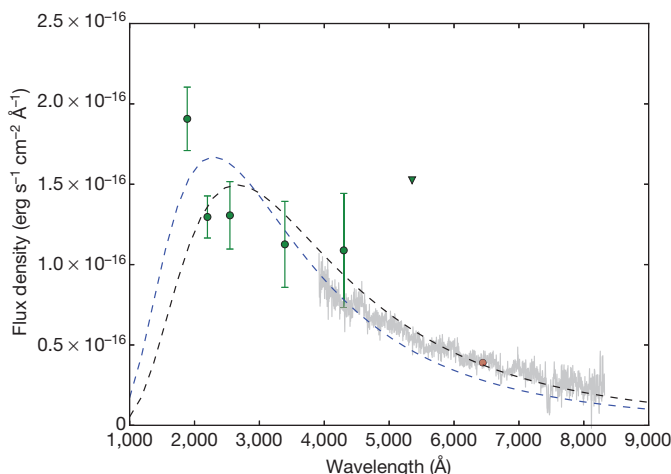


Figure 2 | The spectral energy distribution of iPTF14atg. The spectral energy distribution of iPTF14atg on 2014 May 6 (three days after the explosion) is constructed by using the iPTF r-band magnitude (red), an optical spectrum (grey), and Swift/UVOT measurements (green circles) and upper limit (green triangle). The error bars denote 1σ uncertainties. The blue and black blackbody spectra correspond to the model lightcurves as in Fig. 1.

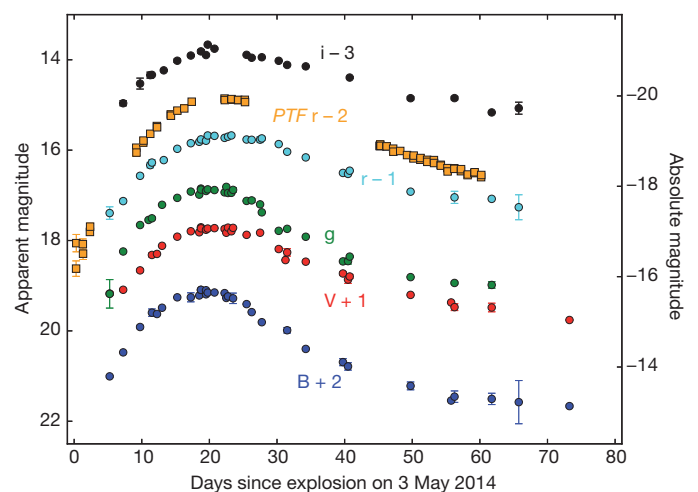


Figure 3 | The multi-colour lightcurve of iPTF14atg. Following the convention, the magnitudes in the B and V bands are in the Vega system while those in the g , r and i bands are in the AB system. Error bars represent 1σ uncertainties. For clearer illustration, the lightcurves in different filters are offset (plus or minus) as indicated by the numbers following the filter labels.

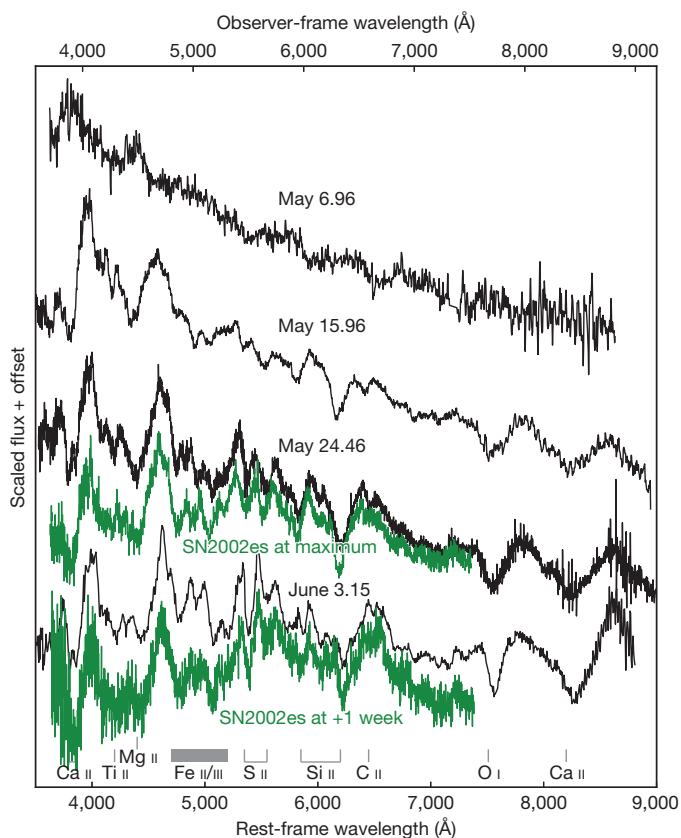


Figure 4 | Spectral evolution of iPTF14atg. The spectra of iPTF14atg (black) are compared with those of SN 2002es (green) at the supernova maximum and a week after maximum. The SN 2002es spectrum at maximum is blue-shifted by $2,000 \text{ km s}^{-1}$ and that at +1 week is blue-shifted by $1,000 \text{ km s}^{-1}$. Ticks at the bottom of the plot label major absorption features in the spectra.

galaxies^{18,20}. Therefore, we tentatively classify iPTF14atg as a high-velocity version of SN 2002es.

Our work, along with recent suggestions for companions in SN 2008ha²¹ and SN 2012Z²², hints that subluminal supernovae with low velocities, such as SN 2002cx and SN 2002es, arise from the single degenerate channel. In contrast, there is mounting evidence that some type Ia supernovae result from the double degenerate channel^{23,24} in which two white dwarfs merge or collide and then explode in a binary or even triple system³. Clearly determining the fraction of type Ia supernovae with companion interaction signatures could disentangle the type Ia supernova progenitor puzzle. Prior to our discovery, searches for companion interaction have been carried out in both ultraviolet^{13,25,26} and ground-based optical data^{27,28}. However, very few supernovae were observed in the ultraviolet within a few days of their explosions. Our observation of iPTF14atg also demonstrates that the interaction signature is not distinctive in the optical bands.

Therefore, rapid ultraviolet follow-up observations of extremely young supernovae or fast-cadence ultraviolet transient surveys are warranted to probe the companion interaction of type Ia supernovae. Given the observed ultraviolet flux of iPTF14atg, ULTRASAT²⁹ (a proposed space telescope aimed at undertaking fast-cadence observations of the ultraviolet sky) should detect such events up to 300 Mpc away from Earth. Factoring in its field-of-view of 210 square degrees, ULTRASAT will probably detect three dozen type Ia supernovae of all kinds within this volume during its two-year mission lifetime. In fact, the ultraviolet flux of the supernova-companion interaction is brighter at earlier phases. Thus, ULTRASAT may discover more such events at greater distances. Since up to a third of type Ia supernovae are sub-luminous¹⁸, the ULTRASAT survey could definitively determine the

fraction of events with companion interaction and thus the rate of events from the single degenerate channel.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 November 2014; accepted 16 March 2015.

- Filippenko, A. V. Optical spectra of supernovae. *Annu. Rev. Astron. Astrophys.* **35**, 309–355 (1997).
- Nugent, P. E. *et al.* Supernova SN 2011fe from an exploding carbon-oxygen white dwarf star. *Nature* **480**, 344–347 (2011).
- Maoz, D., Mannucci, F. & Nelemans, G. Observational clues to the progenitors of type Ia supernovae. *Annu. Rev. Astron. Astrophys.* **52**, 107–170 (2014).
- Riess, A. G. *et al.* Observational evidence from supernovae for an accelerating Universe and a cosmological constant. *Astron. J.* **116**, 1009–1038 (1998).
- Perlmutter, S. *et al.* Measurements of Ω and Λ from 42 high-redshift supernovae. *Astrophys. J.* **517**, 565–586 (1999).
- Sullivan, M. *et al.* SNLS3: constraints on dark energy combining the supernova legacy survey three-year data with other probes. *Astrophys. J.* **737**, 102 (2011).
- Whelan, J. & Iben, I. Jr. Binaries and supernovae of Type I. *Astrophys. J.* **186**, 1007–1014 (1973).
- Wang, B. & Han, Z. Progenitors of type Ia supernovae. *New Astron. Rev.* **56**, 122–141 (2012).
- Kasen, D. Seeing the collision of a supernova with its companion star. *Astrophys. J.* **708**, 1025–1031 (2010).
- Law, N. M. *et al.* The Palomar Transient Factory: system overview, performance, and first results. *Publ. Astron. Soc. Pacif.* **121**, 1395–1408 (2009).
- Gehrels, N. *et al.* The Swift Gamma-Ray Burst Mission. *Astrophys. J.* **611**, 1005–1020 (2004).
- Milne, P. A. *et al.* Near-ultraviolet properties of a large sample of type Ia supernovae as observed with the Swift UVOT. *Astrophys. J.* **721**, 1627–1655 (2010).
- Brown, P. J. *et al.* A Swift look at SN 2011fe: the earliest ultraviolet observations of a type Ia supernova. *Astrophys. J.* **753**, 22 (2012).
- Pan, K.-C., Ricker, P. M. & Taam, R. E. Impact of type Ia supernova ejecta on binary companions in the single-degenerate scenario. *Astrophys. J.* **750**, 151 (2012).
- Yasuda, N. & Fukugita, M. Luminosity functions of type Ia supernovae and their host galaxies from the Sloan Digital Sky Survey. *Astron. J.* **139**, 39–52 (2010).
- Filippenko, A. V. *et al.* The subluminal, spectroscopically peculiar type Ia supernova 1991bg in the elliptical galaxy NGC 4374. *Astron. J.* **104**, 1543–1556 (1992).
- Li, W. *et al.* SN 2002cx: the most peculiar known type Ia supernova. *Publ. Astron. Soc. Pacif.* **115**, 453–473 (2003).
- Foley, R. J. *et al.* Type Iax supernovae: a new class of stellar explosion. *Astrophys. J.* **767**, 57 (2013).
- Ganeshalingam, M. *et al.* The low-velocity, rapidly fading type Ia supernova 2002es. *Astrophys. J.* **751**, 142 (2012).
- White, C. J. *et al.* Slow-speed supernovae from the Palomar Transient Factory: two channels. *Astrophys. J.* **799**, 52 (2015).
- Foley, R. J. *et al.* Possible detection of the stellar donor or remnant for the type Iax supernova 2008ha. *Astrophys. J.* **792**, 29 (2014).
- McCully, C. *et al.* A luminous, blue progenitor system for the type Iax supernova 2012Z. *Nature* **512**, 54–56 (2014).
- Li, W. *et al.* Exclusion of a luminous red giant as a companion star to the progenitor of supernova SN 2011fe. *Nature* **480**, 348–350 (2011).
- González Hernández, J. I. *et al.* No surviving evolved companions of the progenitor of SN 1006. *Nature* **489**, 533–536 (2012).
- Foley, R. J. *et al.* Very early ultraviolet and optical observations of the type Ia supernova 2009ig. *Astrophys. J.* **744**, 38 (2012).
- Brown, P. J. *et al.* Constraints on type Ia supernova progenitor companions from early ultraviolet observations with Swift. *Astrophys. J.* **749**, 18 (2012).
- Hayden, B. T. *et al.* Single or double degenerate progenitors? Searching for shock emission in the SDSS-II type Ia supernovae. *Astrophys. J.* **722**, 1691–1698 (2010).
- Bianco, F. B. *et al.* Constraining type Ia supernovae progenitors from three years of supernova legacy survey data. *Astrophys. J.* **741**, 20 (2011).
- Sagiv, I. *et al.* Science with a wide-field UV transient explorer. *Astron. J.* **147**, 79 (2014).

Acknowledgements We thank A. L. Piro, M. Kromer and J. Cohen for discussions. We also thank A. Waszczak, A. Rubin, O. Yaron, A. De Cia, D. A. Perley, G. E. Duggan, O. Smirnova, S. Papadogiannakis, A. Nyholm, Y. F. Martinez and the staff at the Nordic Optical Telescope and Gemini for observation and data reduction. Some of the data presented here were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and NASA. The observatory was made possible by the generous financial support of the W. M. Keck Foundation. Some data were obtained with the Nordic Optical Telescope, which is operated by the Nordic Optical Telescope Scientific Association at the Observatorio del Roque de los Muchachos, La Palma, Spain. This work also makes use of observations from the Las Cumbres Observatory Global Telescope (LCOGT) network. Research at California Institute of Technology is supported by the National Science Foundation. D.A.H. acknowledges support from the National Science Foundation. A.G.-Y. acknowledges support from the EU/FP7 via an ERC grant, the “Quantum Universe” I-Core programme, the ISF, Minerva and Weizmann-UK grants, and the Kimmel Award. M.M.K. acknowledges generous support from the

Carnegie-Princeton fellowship. Supernova research at the Oskar Klein Centre is supported by the Swedish Research Council and by the Knut and Alice Wallenberg Foundation. The National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy under contract number DE-AC02-05CH11231, provided staff, computational resources, and data storage for this project. The participation of the Los Alamos National Laboratory (LANL) in iPTF is supported by the US Department of Energy as part of the Laboratory Directed Research and Development programme. A portion of this work was carried out at the Jet Propulsion Laboratory under a Research and Technology Development Grant, under contract with the National Aeronautics and Space Administration.

Author Contributions Y.C. initiated the study, conducted analysis and wrote the manuscript. S.R.K. is the iPTF Principal Investigator, and contributed to Swift/UVOT data analysis and manuscript preparation. S.V. and D.A.H. contributed to LCOGT observation, data analysis and manuscript preparation. A.G.-Y. contributed to

manuscript preparation. M.M.K. contributed to Swift, Apache Point Observatory and Gemini-North spectroscopic observations and manuscript preparation. J.J., A.G., J.S., F.T. and R.A. triggered early Nordic Optical Telescope observation and contributed to manuscript preparation. A.H. triggered the Jansky Very Large Array observation and analysed the data. I.S. found the supernova. S.B.C. reduced the P48 data and contributed to manuscript preparation. P.E.N. contributed to manuscript preparation. J.S. and I.A. contributed to building transient vetting and data archiving software. B.D.B., D.I.M., U.D.R. and P.R.W. contributed to the machine learning codes to search for young transients. N.G. is the Swift Principal Investigator.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.C. (ycao@astro.caltech.edu).

METHODS

Discovery. The intermediate Palomar Transient Factory (iPTF) uses the 48-inch Samuel Oschin telescope (P48) at Palomar Observatory, California, USA to characterize optical transients and variable stars¹⁰. A single P48 frame has a field of view of 7.2 square degrees and achieves a detection threshold of $r(\text{AB}) \approx 21$ mag (5 σ , that is, 99.9999% confidence level, CL). From February to June in 2014, iPTF conducted a fast-cadence experiment to search for young transients. Every field was monitored twice, separated by about an hour every night (weather permitting). Transients were identified in real time by a monitoring group aided by machine-learning classifiers^{30–32}. Panchromatic follow-up of young transients was carried out within hours after discovery³³.

The supernova iPTF14atg was discovered on UTC 2014 May 3.29 at $\alpha(\text{J2000}) = 12 \text{ h } 52 \text{ min } 44.8 \text{ s}$ and $\delta(\text{J2000}) = +26^\circ 28' 13''$, about $10''$ east with no measurable offset in declination from the apparent host galaxy IC 831. It had an r -band magnitude of 20.3 upon discovery. No source was detected at the same location on images taken on UT 2014 May 02.25 and 02.29 down to a limiting magnitude of $r \approx 21.4$ (99.9999% CL). No activity had been found at this location in the iPTF archival data in 2013 (3 epochs) and 2014 (101 epochs) down to similar limiting magnitudes. This supernova was also independently discovered by the All-Sky Automated Survey for Supernovae on May 22³⁴ and classified as a SN 1991bg-like type Ia supernova on June 3³⁵.

SDSS (Data Release 12) measured the redshift of IC 831 to be 0.02129³⁶. With the cosmological parameters measured by Planck ($H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.307$, $\Omega_\Lambda = 0.691$ and $\Omega_v = 0.001$)³⁷, we calculate a co-moving distance of 93.7 Mpc and a distance modulus of 34.9 mag for IC 831.

The photometric source data for Fig. 3 is available in the online version of the paper. Both photometric and spectroscopic data will also be made publicly available on WISEREP³⁸ (<http://wiserep.weizmann.ac.il>).

Data acquisition. *Swift* observations and data reduction. Starting on May 6, the Ultraviolet and Optical Telescope (UVOT)³⁹ and the X-ray Telescope (XRT)⁴⁰ onboard the *Swift* Space Observatory⁴¹ observed iPTF14atg for fourteen epochs in May and June (summarized in Extended Data Table 2). To subtract the host galaxy contamination, reference images were taken six months after the supernova explosion. Visual inspection to the reference images ensures that the supernova has faded away.

Photometric measurements of the UVOT images were undertaken with the *uvotsource* routine in the HEASoft package (<http://heasarc.nasa.gov/heasoft/>). Instrumental fluxes of iPTF14atg were extracted with an aperture of radius $3''$ centred at the location determined by the iPTF optical images and the sky background is calculated with an aperture of radius $20''$ in the vicinity of iPTF14atg. The fluxes were then corrected by the growth curves of UVOT point spread functions and for coincidence loss. Then the instrumental fluxes were converted to physical fluxes using the most recent calibration⁴¹. The host galaxy flux is measured with the same aperture in the reference images. The XRT data were analysed with the *Ximage* software in the HEASoft package. We estimated count rate upper limits at a 99.7% CL at the location of iPTF14atg for non-detections.

We use WebPIMMS (<http://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3pimms/w3pimms.pl>) to convert the XRT upper limit on May 6 to physical quantities. As shown in Fig. 2, the optical and ultraviolet data taken on the same day can be approximated by a blackbody model with a temperature $T = 1.1 \times 10^4 \text{ K}$ and radius of 6,000 solar radii. Using the blackbody model with the radius fixed to the above value and setting the interstellar column density of hydrogen to $N_H = 8 \times 10^{19} \text{ cm}^{-2}$ as is appropriate towards this direction⁴², we find that the XRT upper limit of counting rate agrees with $T < 10^5 \text{ K}$.

Ground-based observations and data reduction. As described in Methods subsection ‘Discovery’, P48 observed the iPTF14atg field every night (weather-permitting) until July 2. The host galaxy contamination in these images was removed with the aid of a reference image which was built by stacking twelve P48 pre-supernova images. We then performed the point spread function photometry on the subtracted images. The photometry is calibrated to the PTF-IPAC (Infrared Processing and Analysis Center) catalogue⁴³.

We also triggered LCOGT to follow up iPTF14atg in the griBV filters. Because reference images were not available, we used an image-based model composed of a point spread function and a low-order polynomial to model the supernova light and its underlying galaxy background simultaneously. The photometry is then calibrated to the SDSS catalogue.

Our optical spectroscopic observation log is presented in Extended Data Table 1. Spectroscopic data were reduced with standard routines in the Image Reduction and Analysis Facility (IRAF) and/or Interactive Data Language (IDL).

We also observed iPTF14atg with the Jansky Very Large Array (JVLA) on May 16 at both 6.1 GHz (C-band) and at 22 GHz (K-band). The observation was performed in the A configuration using galaxy J1310+3220 as a phase calibrator

and galaxy 3C286 as a flux calibrator. The data were reduced using standard routines in the CASA software (<http://casa.nrao.edu/>). The observation resulted in a null detection at both bands with an upper limit of $30 \mu\text{Jy}$ (99.7% CL). We note that the radio observation is taken ten days after the discovery of iPTF14atg.

Spherical models for the early ultraviolet pulse. We considered models spherically symmetric to the supernova explosion centre to interpret the early ultraviolet excess seen in iPTF14atg. In the first model, we investigated the possibility that the ultraviolet pulse is powered by radioactive decay. The rise time of a supernova peak is roughly characterized by the diffusive timescale from the radioactive layer to the photosphere, that is

$$\tau \propto \left(\frac{\kappa M_{\text{ej}}}{c v_{\text{exp}}} \right)^{1/2}$$

where M_{ej} is the mass of ejecta outside the radioactive layer, κ is the mean opacity of the ejecta, c is the speed of light in a vacuum and v_{exp} is the mean expansion velocity. To further simplify the situation, we assume that κ remains roughly constant in the rise phase of a supernova. If the mean expansion velocity v_{exp} also does not change considerably, then the ultraviolet pulse of iPTF14atg within four days of its explosion and its main radiation peak observed about 20 days after explosion indicate two distinct radioactive layers. The shallow layer above which the ejecta mass is about 4% of the total ejected mass powered the ultraviolet pulse. Furthermore, if the radioactive element in the shallowest layer is ^{56}Ni , then the ultraviolet luminosity of $3 \times 10^{41} \text{ erg s}^{-1}$ requires a ^{56}Ni mass of 0.01 solar masses. Such a configuration has been widely discussed in double-detonation models where a carbon-oxygen white dwarf accretes mass from a helium star. After the helium shell on the surface of the white dwarf reaches a critical mass, helium burns in a detonation wave with such force that a detonation is also ignited in the interior of the white dwarf, making a supernova explosion of sub-Chandrasekhar mass. However, nuclear burning on the surface not only makes ^{56}Ni , but also generates a layer containing iron-group elements. These elements have a vast number of optically thick lines in the ultraviolet to effectively reprocess ultraviolet photons into longer wavelengths. Therefore we would not have observed the ultraviolet pulse at the early time in this scenario⁴⁴. In addition, some double-detonation models predict weak Si II 6355 Å absorption near the supernova peak⁴⁵ which contradicts the observed deep Si II absorption in iPTF14atg. Therefore we consider this model as not consistent with observations.

A second model is that the emission arises from circumstellar gas around the progenitor. The circumstellar gas is heated up by either high-energy photons from the supernova shock breakout (case 1) or the supernova shock itself (case 2). We assume that the circumstellar gas is optically thin and thus the plausible radiation mechanism is bremsstrahlung. Consider a simple model of a sphere of radius R_s that contains pure material with an atomic number Z and a mass number A . The material is completely ionized, so the electron density n_e and the ion density n_i are related by $n_e = Zn_i$. The bremsstrahlung luminosity is

$$L_{\text{ff}} = 1.4 \times 10^{-27} T^{1/2} n_e n_i Z^2 \times \frac{4\pi}{3} R_s^3$$

where all physical quantities are in centimetre–gram–second (CGS) units. We further assume the critical case that the optical depth of the sphere is unity, that is

$$\tau_s = n_e \sigma_T R_s = 1$$

Then we can derive analytical expressions for the luminosity, the total mass of the sphere M and the thermal energy Q in terms of A , Z , R_s and temperature T , that is

$$L_{\text{ff}} = 5.0 \times 10^{41} Z R_{17} T_5^{1/2} \text{ erg s}^{-1}$$

$$M = \frac{4\pi}{3} R_s^3 \times n_i A u = 53 \frac{A}{Z} R_{17}^2 \text{ solar masses}$$

$$Q = \frac{4\pi}{3} R_s^3 \times (n_i + n_e) \times \frac{3}{2} k_B T = 1.3 \times 10^{48} \times \frac{Z+1}{Z} \times R_{17}^2 T_5 \text{ erg}$$

where $R_s = R_{17} \times 10^{17} \text{ cm}$, $T = T_5 \times 10^5 \text{ K}$, and u is the atomic mass unit. Because no hydrogen is seen in the iPTF14atg spectra, we assume that the sphere is dominated by helium, so that $A = 4$ and $Z = 2$.

In case 1, the circumstellar gas is heated up by the high-energy photons from the supernova shock breakout. The temperature of the gas is roughly 11,000 K, as determined by the optical–ultraviolet spectral energy distribution. To account for a ultraviolet luminosity of $3 \times 10^{41} \text{ erg s}^{-1}$, the radius R_s has to be as large as $3 \times 10^{16} \text{ cm}$. The total mass of the sphere would be ten solar masses and the total thermal energy would be 10^{47} erg . If the optical depth of the sphere is larger than unity, then we will end up with an even more massive sphere. So we are forced to invoke a sphere containing a mass much larger than a typical type Ia supernova. The absence of strong Na I D lines also argue against such massive circumstellar

material. In addition, the elliptical host galaxy with no star-forming activity also excludes the existence of massive stars.

In case 2, the circumstellar gas is ionized by the supernova shock. The supernova shock has a typical velocity between $20,000 \text{ km s}^{-1}$ and $5,000 \text{ km s}^{-1}$. Hence, within four days of the supernova explosion, the supernova shock travelled to $R_s \approx 10^{15} \text{ cm}$. To account for the ultraviolet pulse, this small radius then requires an extremely high temperature of 10^7 K , which is inconsistent with the observed spectral energy distribution. Therefore we discard this model.

Supernova specification. We performed comparative analysis among iPTF14atg, SN 1991bg, SN 2002cx and SN 2002es on the photometric and spectroscopic evolution and host galaxy and demonstrate that iPTF14atg is likely to belong to the SN 2002es family.

Photometry. The multiband lightcurve of iPTF14atg is shown in Fig. 3. Note that there is an approximately 0.2 mag difference between the PTF *r*-band and LCOGT *r*-band magnitudes. We calculated synthetic photometry using the iPTF14atg spectra and the filter transmission curves and found that this difference was mainly due to the filter difference.

Because iPTF14atg is not a normal type Ia supernova, the usual lightcurve fitting tools for normal type Ia supernovae (for example, SALT2⁴⁶, SNooPy²⁴⁷) are not suitable to determine the lightcurve features. Thus we fit a 5th-order polynomial to the *B*-band lightcurve and derived a *B*-band peak magnitude of 17.1 mag on May 22.15 and $\Delta m_{15} = 1.2 \text{ mag}$ (Δm_{15} is the magnitude change of a type Ia supernova in 15 days from its peak, which astronomers use to measure the shape of its light curve). We also infer that the line-of-sight extinction is low because the Galactic extinction in this direction is $A_B = 0.032$ and because we do not see any sign of strong Na I D absorption in all of our low-resolution and medium-resolution spectra of iPTF14atg. Hence, given the host galaxy distance modulus of 34.9 mag, we conclude that iPTF14atg has an absolute peak magnitude of -17.8 mag and that iPTF14atg is a subluminal outlier of the well-established relation between the peak magnitude and Δm_{15} (ref. 48).

We compare iPTF14atg with the three major families of subluminal type Ia supernovae with the prototypical events SN 1991bg, SN 2002cx and SN 2002es. From Extended Data Fig. 1, we can see that: (1) the peak magnitude of iPTF14atg is brighter than that of SN 1991bg, similar to SN 2005hk (a typical SN 2002cx-family event), and fainter than SN 2002es. However, both SN 2002cx and SN 2002es families have large ranges of peak magnitudes^{18,20}. (2) iPTF14atg evolves more slowly than SN 1991bg in both rise and decline phases. (3) iPTF14atg has a slower rise than SN 2005hk. (4) Unlike SN 2002es, iPTF14atg does not have a break in the lightcurve about 30 days after the peak. A caveat about this comparison is that the lightcurve of iPTF14atg, especially the very early part, might be distorted by the supernova-companion collision.

We present the near-ultraviolet and optical colour evolution of iPTF14atg in Extended Data Fig. 2 and compare it with SN 2011fe (also known as PTF11kly, a normal type Ia supernova in the Pinwheel Galaxy 6 Mpc away from Earth)^{13,49}, SN 2002es, SN 2005hk and SN 1991bg. The figure shows that iPTF14atg was initially bluer in $uvm2 - uvw1$ by more than two magnitudes than was SN 2011fe, which is classified as a near-ultraviolet blue event¹³. Though SN 2011fe gradually becomes bluer while approaching its peak, iPTF14atg remains the same colour and still bluer at peak than SN 2011fe by one magnitude. The optical colour, indicated by $B - V$, of iPTF14atg was initially red; it quickly became blue within a few days and then followed the evolution of SN 2002es. Though SN 2011fe was also red initially, it gradually became blue during the supernova rise, reached its bluest colour near the supernova peak and turned red later on.

Spectroscopy. The spectral evolution of iPTF14atg is presented in Extended Data Fig. 3 and is compared with those of SN 1991bg, SN 2005hk and SN 2002es in Extended Data Fig. 4. On May 6 (within four days after the explosion) when the ultraviolet excess was detected, the spectrum of iPTF14atg consisted of a blue continuum superposed by some weak and broad absorption features. In Fig. 2 and Extended Data Fig. 3 we tentatively identified Si II, S II and Ca II lines. Combining the photometry from Swift/UVOT, the spectral energy distribution can be approximated by a blackbody spectrum of temperature 11,000 K and radius 6,000 solar radii (Fig. 2). None of the known subluminal type Ia supernovae have been observed at such an early time and are therefore unavailable for comparison, so we turned to SN 2011fe. Unlike iPTF14atg, the spectra of SN 2011fe taken within two days after explosion show clear absorption features commonly seen in a pre-maximum type Ia supernova, such as Si II, S II, Mg II, O I and Ca II. We therefore suggest that this spectrum of iPTF14atg has a dominant thermal component from the supernova-companion interaction and a weak supernova component from intact regions of the supernova photosphere. In the next spectrum taken three days later, spectral features such as Si II and Ca II have emerged. In the spectrum taken on May 11, we clearly identified Si II around 6,100 Å, with its minimum at a velocity of $10,000 \text{ km s}^{-1}$. This velocity is lower than that of a normal type Ia supernova at a similar phase⁵⁰.

In the spectrum taken on May 15 (about a week before its maximum brightness; see the first panel of Extended Data Fig. 4), we identified absorption features such as Si II, the S II 'W' around 5,000 Å and O I and concluded that iPTF14atg is a type Ia supernova based on the presence of Si II and S II. The whole spectrum matched well to the SN 1991bg spectrum at a similar phase after we redshifted the SN 1991bg spectrum by $3,000 \text{ km s}^{-1}$. The difference was that iPTF14atg did not have a Ti trough around 4,200 Å as deep as that of SN 1991bg. This iPTF14atg spectrum showed little similarity to that of SN 2005hk, a typical SN 2002cx-like event.

Near the supernova peak (see the second panel of Extended Data Fig. 4), the spectrum of iPTF14atg shared similar absorption features with all three families of subluminal type Ia supernovae, though the depth of the absorptions and the continuum shapes differed among them. iPTF14atg had a velocity lower than SN 1991bg but higher than SN 2005hk. The best match to the overall spectral shape was between iPTF14atg and SN 2002es.

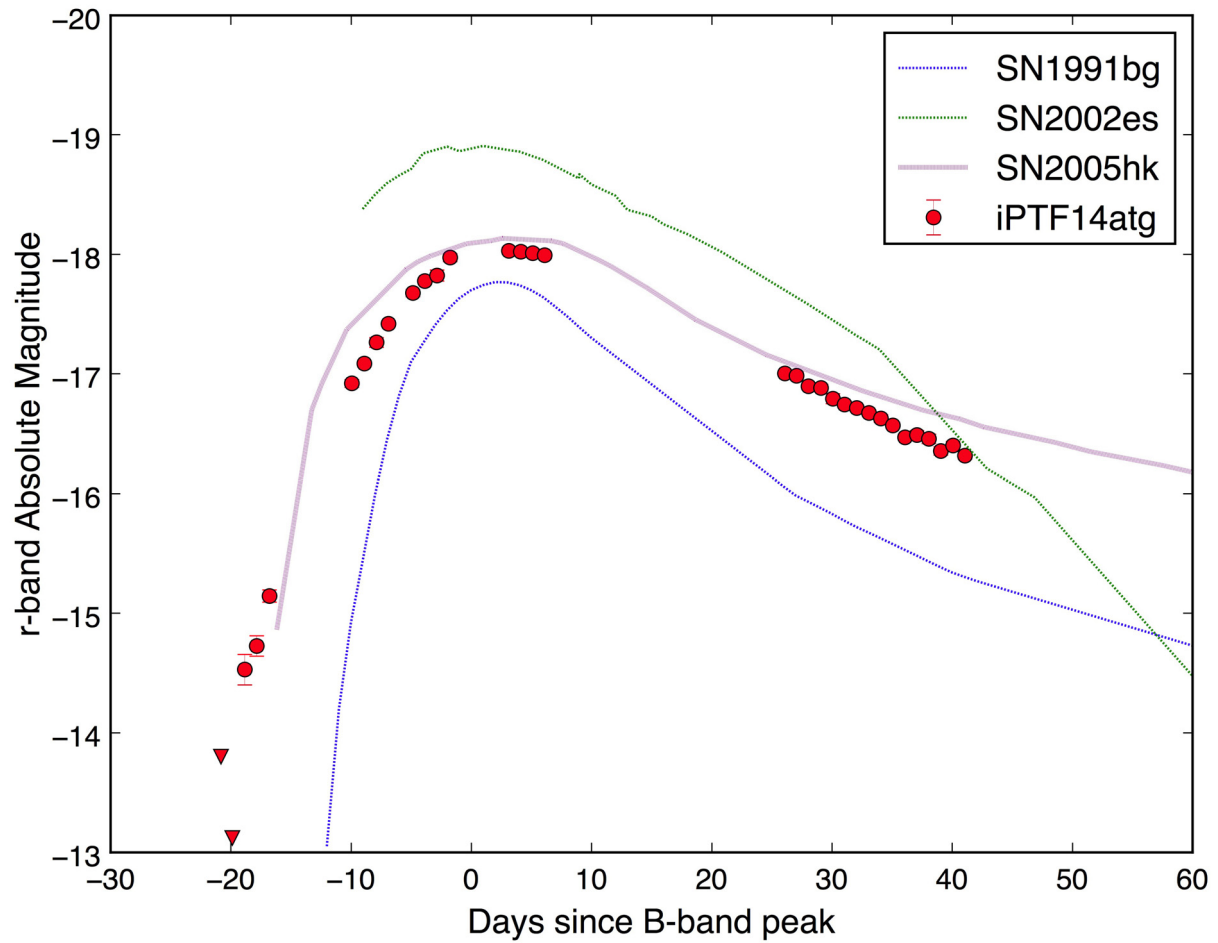
The post-maximum spectral evolution (see the third to fifth panels of Extended Data Fig. 4) shows that iPTF14atg shares many spectral similarities with SN 2002cx-like events, but differences in the near infrared part of the late-time spectrum taken two months after the supernova peak (see the fifth panel of Extended Data Fig. 4) disfavoured its classification of a SN 2002cx-family. In contrast, iPTF14atg spectra matched well with the limited spectral information available for SN 2002es-like events.

Another interesting feature of iPTF14atg is strong and persistent C II 6,580 Å absorption. The absorption feature C II, is sometimes seen in pre-maximum spectra of normal type Ia supernovae^{51,52} and always disappears before the supernova peak. Only in very few cases^{50,53} did the carbon absorption feature exist at a velocity decreasing from $11,000 \text{ km s}^{-1}$ to $9,000 \text{ km s}^{-1}$ after maximum light. It also has been reported in pre-max or max spectra of SN 2002cx-like events but not in post-maximum spectra^{54,55}. SN 2002es may show very weak C II three days after maximum¹⁹. In the case of iPTF14atg, the C II 6,580 Å absorption is first seen at a velocity of $11,000 \text{ km s}^{-1}$ in the spectrum taken on May 11 (about 12 days before maximum light). Its velocity decreased to $6,000 \text{ km s}^{-1}$ near the maximum light. After the maximum light, its velocity kept decreasing. It was detected at a velocity of $4,000 \text{ km s}^{-1}$ in the spectrum taken on June 6 (two weeks after maximum light) but not later. The long-lasting carbon feature indicates that both high-velocity shallow layers and low-velocity deep layers in the iPTF14atg explosion are carbon-rich, which is evidence for incomplete burning extending deep into the ejecta. The incomplete burning is consistent with a pure deflagration⁵⁶.

Host galaxy. The host galaxy IC 831 is morphologically classified as an elliptical galaxy⁵⁷ or an S0 galaxy⁵⁸. Its SDSS spectrum shows very weak H α and [O III] emission. This suggests that the host galaxy has little star-forming activity. iPTF14atg occurred about 4.6 kiloparsecs from the centre of IC 831. In all iPTF14atg spectra, we did not detect any H α emission either at the supernova location or underlying in the galaxy background, suggesting that iPTF14atg was born in an old population. This strongly argues against that iPTF14atg is a core-collapse supernova. Furthermore, SN 2002cx-like events prefer star-forming regions while SN 2002es-like events are all found in the passive galaxies^{18,20}. Hence the nature of its host galaxy makes iPTF14atg more likely to belong to the SN 2002es family.

30. Brink, H. *et al.* Using machine learning for discovery in synoptic survey imaging data. *Mon. Not. Astron. R. Soc. Lond.* **435**, 1047–1060 (2013).
31. Woźniak, P. R. *et al.* Automated variability selection in time-domain imaging surveys using sparse representations with learned dictionaries. *Am. Astron. Soc. Meet. Abstr.* **221**, 431.05 (2013).
32. Bue, B. D., Wagstaff, K. L., Rebbapragada, U. D., Thompson, D. R. & Tang, B. Astronomical data triage for rapid science return. In *Proc. 2014 Conf. on 'Big Data from Space (BIDS'14)'* 206 (2014).
33. Gal-Yam, A. *et al.* Real-time detection and rapid multiwavelength follow-up observations of a highly subluminal type II-P supernova from the Palomar Transient Factory Survey. *Astrophys. J.* **736**, 159 (2011).
34. Holoiu, T. W.-S. *et al.* ASAS-SN discoveries of a probable supernova in IC 0831 and a possible extreme (delta $V > 6.6 \text{ mag}$) M-dwarf flare. *Astron. Teleg.* **6168**, 1 (2014).
35. Wagner, R. M. *et al.* Spectroscopic classification of ASASSN-14bd. *Astron. Teleg.* **6203**, 1 (2014).
36. Bolton, A. S. *et al.* Spectral classification and redshift measurement for the SDSS-III baryon oscillation spectroscopic survey. *Astron. J.* **144**, 144 (2012).
37. Planck Collaboration *et al.* Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).
38. Yaron, O. & Gal-Yam, A. WiSeREP—an interactive supernova data repository. *Publ. Astron. Soc. Pacif.* **124**, 668–681 (2012).
39. Roming, P. W. A. *et al.* in *X-Ray and Gamma-Ray Instrumentation for Astronomy XIII* (eds Flanagan, K. A. & Siegmund, O. H. W.) *Proc. SPIE Conf. Ser.* **5165**, 262–276 (2004).
40. Burrows, D. N. *et al.* in *X-Ray and Gamma-Ray Instrumentation for Astronomy XIII* (eds Flanagan, K. A. & Siegmund, O. H. W.) *Proc. SPIE Conf. Ser.* **5165**, 201–216 (2004).
41. Breeveld, A. A. *et al.* An updated ultraviolet calibration for the Swift/UVOT. *AIP Conf. Ser.* **1358**, 373–376 (2011).

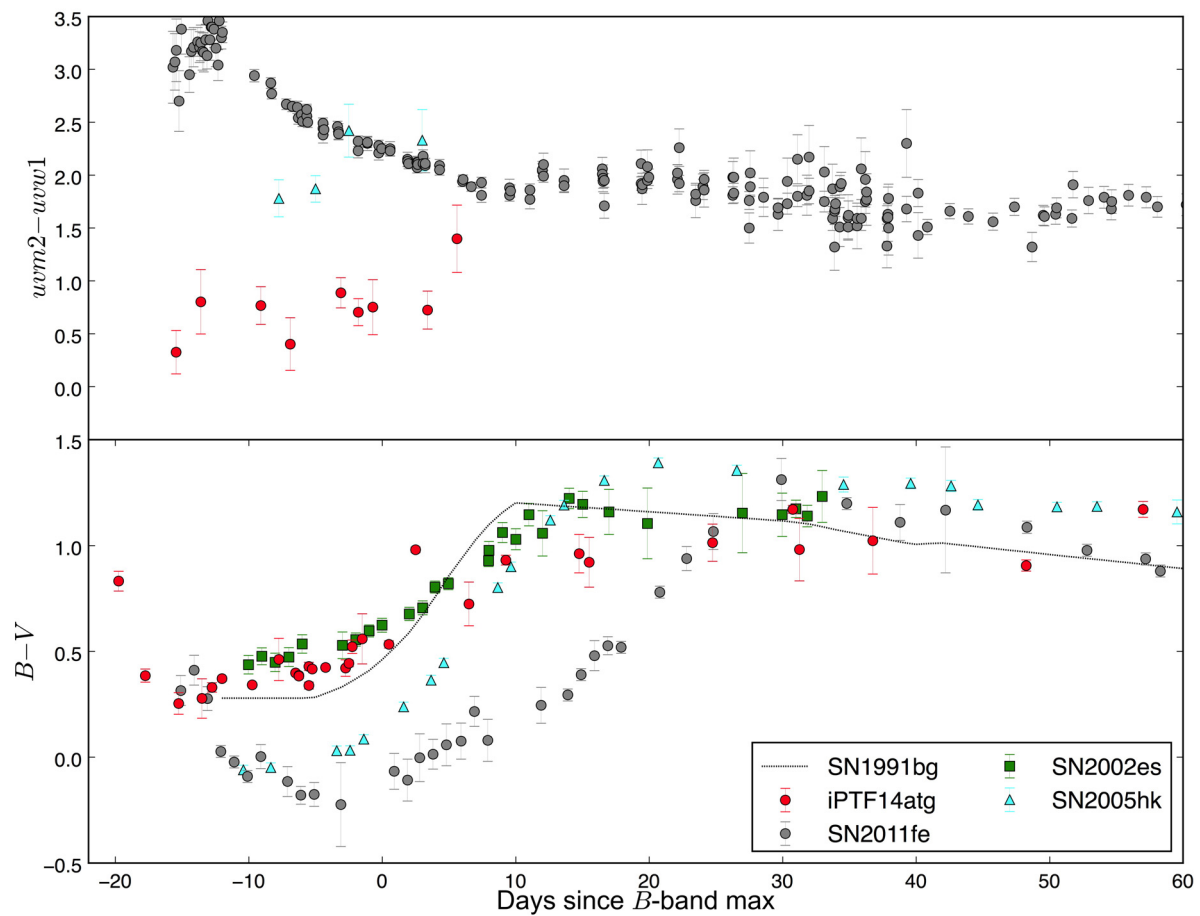
42. Kalberla, P. M. W. *et al.* The Leiden/Argentine/Bonn (LAB) survey of galactic HI. Final data release of the combined LDS and IAR surveys with improved stray-radiation corrections. *Astron. Astrophys.* **440**, 775–782 (2005).
43. Ofek, E. O. *et al.* The Palomar Transient Factory photometric calibration. *Publ. Astron. Soc. Pacif.* **124**, 62–73 (2012).
44. Kromer, M. *et al.* Double-detonation sub-Chandrasekhar supernovae: synthetic observables for minimum helium shell mass models. *Astrophys. J.* **719**, 1067–1082 (2010).
45. Nugent, P., Baron, E., Branch, D., Fisher, A. & Hauschildt, P. H. Synthetic spectra of hydrodynamic models of type Ia supernovae. *Astrophys. J.* **485**, 812–819 (1997).
46. Guy, J. *et al.* SALT2: using distant supernovae to improve the use of type Ia supernovae as distance indicators. *Astron. Astrophys.* **466**, 11–21 (2007).
47. Burns, C. R. *et al.* The Carnegie Supernova Project: light-curve fitting with SNooPy. *Astron. J.* **141**, 19 (2011).
48. Phillips, M. M. The absolute magnitudes of type Ia supernovae. *Astrophys. J.* **413**, L105–L108 (1993).
49. Vinkó, J. *et al.* Testing supernovae Ia distance measurement methods with SN 2011fe. *Astron. Astrophys.* **546**, A12 (2012).
50. Parrent, J. T. *et al.* Analysis of the early-time optical spectra of SN 2011fe in M101. *Astrophys. J.* **752**, L26 (2012).
51. Thomas, R. C. *et al.* Type Ia supernova carbon footprints. *Astrophys. J.* **743**, 27 (2011).
52. Silverman, J. M. & Filippenko, A. V. Berkeley supernova Ia program—IV. Carbon detection in early-time optical spectra of type Ia supernovae. *Mon. Not. Astron. R. Soc. Lond.* **425**, 1917–1933 (2012).
53. Cartier, R. *et al.* Persistent C II absorption in the normal type Ia supernova 2002fk. *Astrophys. J.* **789**, 89 (2014).
54. Foley, R. J. *et al.* Early- and late-time observations of SN 2008ha: additional constraints for the progenitor and explosion. *Astrophys. J.* **708**, L61–L65 (2010).
55. Parrent, J. T. *et al.* A study of carbon features in type Ia supernova spectra. *Astrophys. J.* **732**, 30 (2011).
56. Gamezo, V. N., Khokhlov, A. M., Oran, E. S., Chtchelkanova, A. Y. & Rosenberg, R. O. Thermonuclear supernovae: simulations of the deflagration stage and their implications. *Science* **299**, 77–81 (2003).
57. Scodreggio, M., Giovanelli, R. & Haynes, M. P. The universality of the fundamental plane of E and SO galaxies: sample definition and i-band photometric data. *Astron. J.* **116**, 2728–2737 (1998).
58. Huchra, J. P. *et al.* The 2MASS Redshift Survey—description and data release. *Astrophys. J.* **199** (Suppl.), 26 (2012).
59. Oke, J. B. & Gunn, J. E. An efficient low resolution and moderate resolution spectrograph for the Hale telescope. *Publ. Astron. Soc. Pacif.* **94**, 586 (1982).
60. Faber, S. M. *et al.* in *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes* (eds Iye, M. & Moorwood, A. F. M.) *Proc. SPIE Conf. Ser.* **4841**, 1657–1669 (2003).
61. Oke, J. B. *et al.* The Keck low-resolution imaging spectrometer. *Publ. Astron. Soc. Pacif.* **107**, 375 (1995).
62. Hook, I. M. *et al.* The Gemini-North Multi-Object Spectrograph: performance in imaging, long-slit, and multi-object spectroscopic modes. *Publ. Astron. Soc. Pacif.* **116**, 425–440 (2004).



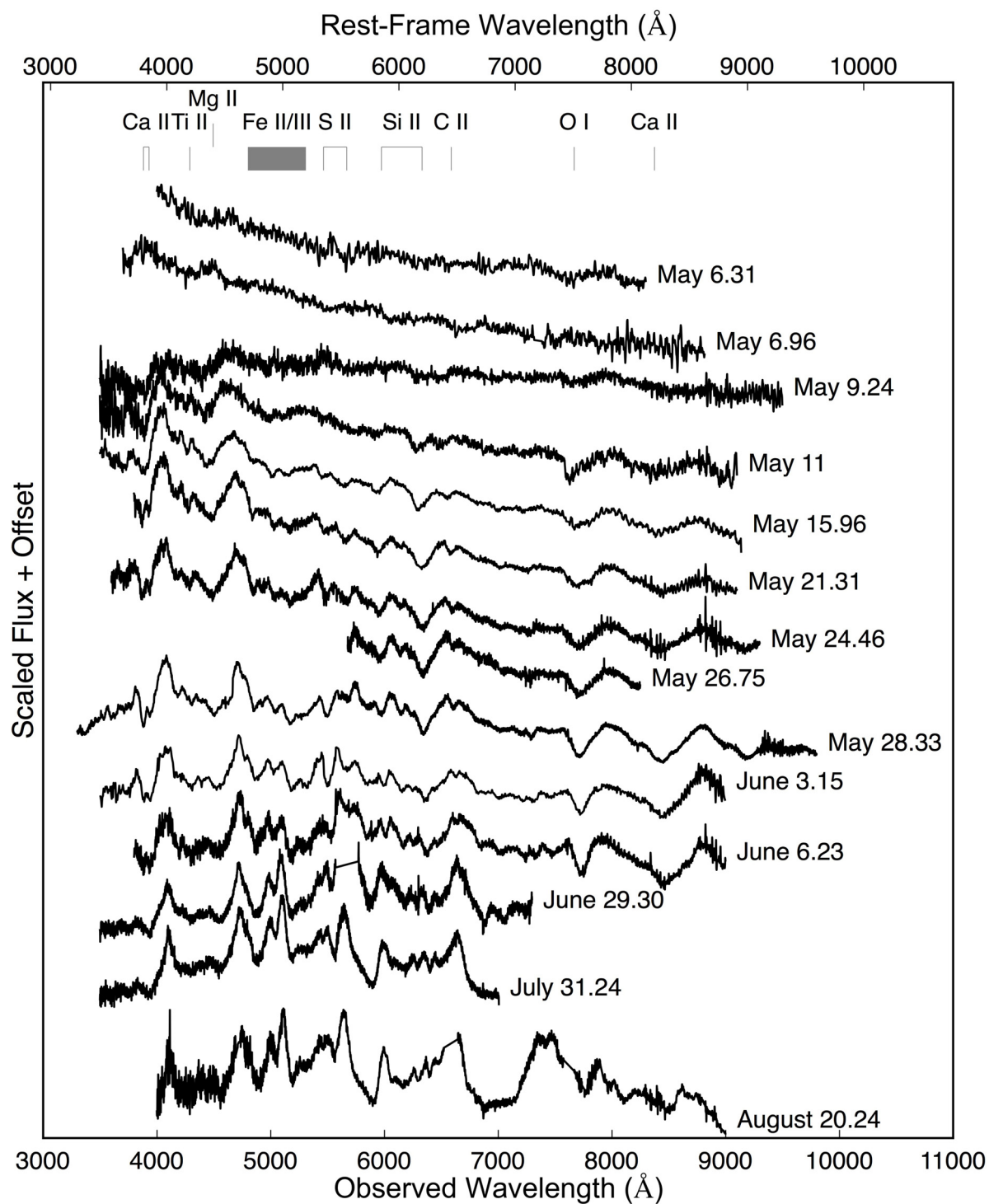
Extended Data Figure 1 | Comparative analysis of iPTF14atg lightcurve.

The lightcurves of iPTF14atg are compared to the Nugent template light curves of SN 1991bg-like events, (the Nugent supernova template is available at https://c3.lbl.gov/nugent/nugent_templates.html), and observed lightcurves of

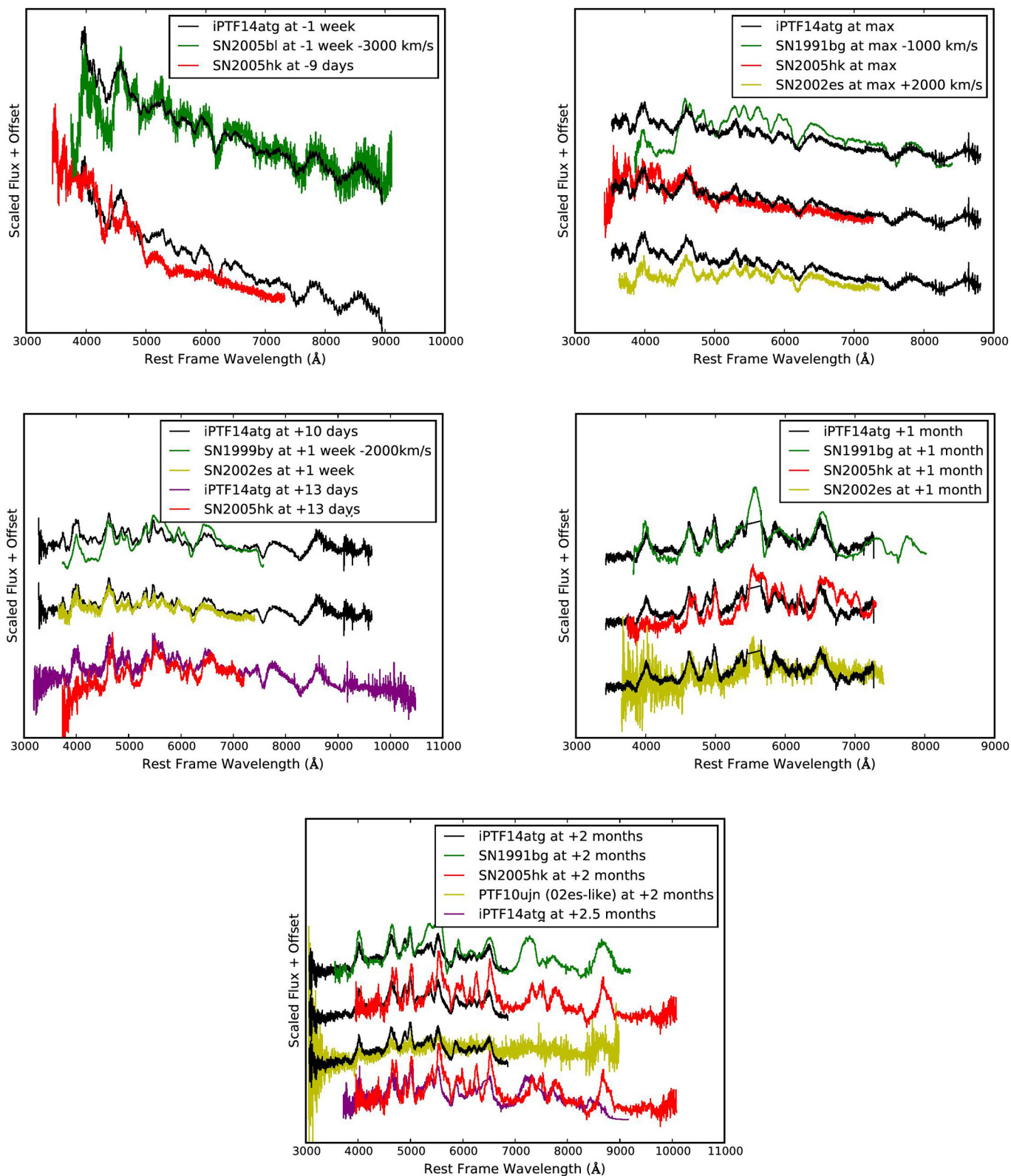
a typical SN 2002cx-like event SN 2005hk and SN 2002es. The error bars denote 1σ uncertainties of magnitudes. The red triangles are upper limits at a 99.9999% CL for non-detections of iPTF14atg.



Extended Data Figure 2 | Comparative analysis of iPTF14atg colour evolution. The colour curves of iPTF14atg are compared to SN 1991bg, SN 2005hk, SN 2002es and a normal event SN 2011fe. The error bars denote 1σ uncertainties.



Extended Data Figure 3 | The spectral evolution of iPTF14atg. Ticks at the top of the figure label major absorption features.



Extended Data Figure 4 | Comparative analysis of iPTF14atg spectra. The spectra of iPTF14atg at different phases are compared with those of SN 1991bg, SN 2005bl (SN 1991bg-like), SN 2005hk, SN 2002es and PTF10ujn (SN 2002es-like) at similar phases.

Extended Data Table 1 | Spectroscopic observation log

Date (UT)	Telescope/Instrument	$\Delta\lambda$ (Å)	Wavelength (Å)	Observer	Data Reducer
May 6.32	ARC-3.5m/DIS ^a	10	3500 - 9500	Cao	Cao
May 6.96	NOT/ALFOSC ^b	16.2	3500 - 9000	O. Smirnova	Taddia
May 9.25	ARC-3.5m/DIS ^a	10	3500 - 9500	Kasliwal	Cao
May 11.04	NOT/ALFOSC ^b	16.2	3500 - 9000	Y. F. Martinez	Taddia
May 15.96	NOT/ALFOSC ^b	16.2	3500 - 9000	A. Nyholm	S. Papadogiannakis
May 21.31	ARC-3.5m/DIS ^a	10	3500 - 9500	Cao	Cao
May 24.21	Hale/DBSP ^c	10	3300 - 10000	A. Waszczak	A. Rubin & O. Yaron
May 26.35	Keck-II/DEIMOS ^d	1.5	5700 - 8200	Cao	A. De Cia
May 28.33	Keck-I/LRIS ^e	7	3300 - 10000	D. A. Perley	D. A. Perley
June 3.15	ARC-3.5m/DIS ^a	10	3500 - 9500	Cao	Cao
June 6.23	Hale/DBSP ^c	10	3300 - 10000	A. Waszczak	O. Yaron
June 29.30	Keck-I/LRIS ^e	7	3300 - 5500	Cao & G. E. Duggan	D. A. Perley
		4.7	5800 - 7400		
July 30.24	Keck-I/LRIS ^e	4	3300 - 5500	Cao	D. A. Perley
		2.5	5400 - 7000		
August 20.24	Gemini-N/GMOS ^f	3	4000 - 9000		Kasliwal

^a The Dual Image Spectrograph (DIS) on the ARC 3.5 m telescope at the Apache Observatory, New Mexico, USA.

^b The Andalucia Faint Object Spectrograph and Camera (ALFOSC) on the Nordic Optical Telescope (NOT) at La Palma, Spain.

^c The Double Spectrograph (DBSP)⁵⁹ on the Palomar 200-inch Hale telescope at Palomar Observatory, California, USA.

^d The DEep Imaging Multi-Object Spectrograph (DEIMOS)⁶⁰ on the Keck-II telescope at Mauna Kea, Hawaii, USA.

^e The Low Resolution Imaging Spectrometer (LRIS)⁶¹ on the Keck-I telescope at Mauna Kea, Hawaii, USA.

^f The Gemini Multi-Object Spectrograph (GMOS)⁶² on the Gemini-N telescope at Mauna Kea, Hawaii, USA.

Extended Data Table 2 | Swift Observation of iPTF14atg

UT Time	UVOT (counts/sec) ^a						XRT (counts/sec) ^b
	uvw2	uvm2	uvw1	u	b	v	
May 06.67 - 06.74	0.297 ± 0.028	0.176 ± 0.014	0.399 ± 0.044	1.251 ± 0.116	2.354 ± 0.168	1.491 ± 0.131	< 3.7 × 10 ⁻³
May 08.53 - 08.61	0.112 ± 0.019	0.096 ± 0.014	0.324 ± 0.041	1.374 ± 0.151	2.444 ± 0.212	1.472 ± 0.163	< 4.9 × 10 ⁻³
May 12.98 - 13.12	0.208 ± 0.019	0.182 ± 0.019	0.578 ± 0.044	2.971 ± 0.152	4.515 ± 0.194	2.288 ± 0.136	< 5.2 × 10 ⁻³
May 15.25 - 15.38	0.351 ± 0.057	0.296 ± 0.041	0.681 ± 0.100	3.636 ± 0.414	5.557 ± 0.531	2.945 ± 0.371	< 2.2 × 10 ⁻³
May 18.99 - 19.05	0.452 ± 0.051	0.297 ± 0.024	1.081 ± 0.094	5.381 ± 0.372	6.600 ± 0.417	3.580 ± 0.285	< 8.8 × 10 ⁻³
May 20.31 - 20.46	0.404 ± 0.034	0.325 ± 0.024	0.956 ± 0.067	5.758 ± 0.307	7.084 ± 0.347	3.178 ± 0.214	< 6.5 × 10 ⁻³
May 21.45 - 21.46	0.437 ± 0.067	0.262 ± 0.042	0.809 ± 0.115	6.015 ± 0.578	7.454 ± 0.669	3.178 ± 0.407	< 2.4 × 10 ⁻²
May 25.45 - 25.65	0.166 ± 0.019	0.156 ± 0.014	0.483 ± 0.041	4.228 ± 0.219	6.808 ± 0.291	3.590 ± 0.195	< 3.6 × 10 ⁻³
May 27.65 - 27.85	0.124 ± 0.023	0.084 ± 0.014	0.440 ± 0.053	3.595 ± 0.274	6.804 ± 0.400	3.494 ± 0.266	< 6.8 × 10 ⁻³
May 30.24 - 30.52	0.078 ± 0.018	0.029 ± 0.009	0.352 ± 0.045	2.551 ± 0.227	4.883 ± 0.330	3.490 ± 0.257	< 7.0 × 10 ⁻³
Jun 07.38 - 07.65	0.041 ± 0.011	0.019 ± 0.006	0.187 ± 0.028	1.101 ± 0.119	2.936 ± 0.195	2.279 ± 0.162	< 3.8 × 10 ⁻³
Jun 17.39 - 17.60	0.052 ± 0.014	0.027 ± 0.009	0.143 ± 0.023	0.859 ± 0.102	2.235 ± 0.168	1.782 ± 0.196	< 5.4 × 10 ⁻³
Jun 21.37 - 21.45	0.039 ± 0.011	0.023 ± 0.006	0.127 ± 0.024	0.833 ± 0.107	2.377 ± 0.176	1.851 ± 0.143	< 4.9 × 10 ⁻³
Jun 25.24 - 25.53	0.045 ± 0.012	0.018 ± 0.007	0.120 ± 0.026	0.801 ± 0.113	1.624 ± 0.153	0.413 ± 0.039	< 4.6 × 10 ⁻³
Nov 12.04 - 12.11 ^c	0.004 ± 0.010	0.017 ± 0.008	0.080 ± 0.025	0.559 ± 0.115	1.576 ± 0.189	0.953 ± 0.156	

^a The uncertainties are at a 68.3% CL.
^b The upper limits are at a 99.7% CL.
^c This is a reference epoch to remove host galaxy contamination.

No signature of ejecta interaction with a stellar companion in three type Ia supernovae

Rob P. Olling¹, Richard Mushotzky¹, Edward J. Shaya¹, Armin Rest², Peter M. Garnavich³, Brad E. Tucker^{4,5}, Daniel Kasen^{5,6}, Steve Margheim⁷ & Alexei V. Filippenko⁵

Type Ia supernovae are thought to be the result of a thermonuclear runaway in carbon/oxygen white dwarfs, but it is uncertain whether the explosion is triggered by accretion from a non-degenerate companion star or by a merger with another white dwarf. Observations of a supernova immediately following the explosion provide unique information on the distribution of ejected material¹ and the progenitor system. Models predict² that the interaction of supernova ejecta with a companion star or circumstellar debris lead to a sudden brightening lasting from hours to days. Here we present data for three supernovae that are likely to be type Ia observed during the Kepler mission³ with a time resolution of 30 minutes. We find no signatures of the supernova ejecta interacting with nearby companions. The lack of observable interaction signatures is consistent with the idea that these three supernovae resulted from the merger of binary white dwarfs or other compact stars such as helium stars.

Barring extraordinary luck, the continuous monitoring of many galaxies is required to observe supernovae immediately after ignition. Our Kepler programme monitored 400 galaxies for two to three years, discovering five supernovae near explosion. Previously only a handful of type Ia supernovae (SN 2009ig⁴, SN 2010jn⁵, SN 2011fe^{6,7}, SN 2012cg⁸, SN 2013dy⁹, and SN 2014j¹⁰) have been observed during the first few days after the explosion. The second Sloan Digital Sky Survey (SDSS-II) contains one of the largest samples of early type Ia

supernova lightcurves¹¹, but that survey had an average cadence exceeding four days. No clear companion interaction signature was found in the SDSS-II Supernova Survey¹², thus ruling out red giants or larger companions. Similar null results were found in analyses of 87 supernovae from the Supernova Legacy Survey (SNLS) survey¹³, 61 supernovae from the Lick Observatory Supernova Search (LOSS) survey¹⁴, and a set of about 700 lightcurves from various sources¹⁵.

Before our Kepler observations, the earliest supernova observations were for SN 2013dy (discovered an estimated 2.4 h post-explosion)⁹ and SN 2011fe (11 h post-explosion)⁷. The determination of the explosion time (actually, the time of first light), however, depends strongly upon the model used to fit the lightcurve⁹. Our analysis of the Kepler supernovae suggests that to determine (in a model-independent way) the explosion time directly to an accuracy of better than 7 h, high-quality data taken at very high cadence are needed.

Analysis of the Kepler data is complicated by a variety of systematic effects on long timescales. Because our procedure co-adds many more pixels centred on a target than the Kepler project^{3,16} pipeline, we account for more light from the source and our results are less sensitive to various systematic effects such as centroid motion or variations in the point-spread function. We also implemented methods that eliminate diffuse background emission in the Kepler data, greatly improving the long-term photometric stability. After correcting for these effects, the lightcurves show variance consistent with Poisson statistics.

Table 1 | Properties of galaxies with supernovae discovered with Kepler

	Kepler supernova ID	KSN 2012a	KSN 2011b	KSN 2011c
1	Kepler ID	8957091	3111451	7889229
2	Right ascension (J2000)	19 h 33 min 30.10 s	19 h 20 min 37.50 s	19 h 24 min 46.10 s
3	Declination (J2000)	45° 15' 01''	38° 15' 08''	43° 40' 51''
4	Kepler magnitude	17.61	15.93	16.87
5	Redshift	0.086	0.052	0.144
6	Extinction (mag)	0.42	0.40	0.37
7	Supernova magnitude at peak	19.25	18.01	20.61
8	Supernova absolute magnitude at peak	−19.14	−19.07	−17.73
9	$\Delta(\text{MLCS2k2})$	0.66 ± 0.11	0.18 ± 0.06	0.84 ± 0.26
10	$C \times 1,000$	5.98 ± 2.02	2.21 ± 0.08	0.71 ± 0.64
11	α	2.12 ± 0.14	2.44 ± 0.15	2.58 ± 0.33
12	$t_{\text{first light}}$ (days)	−15.70, −0.29, +0.30	−18.11, −0.40, +0.30	−20.0, −2.1, +1.5
13	$t_{50\%,\text{before}}$ (days)	−7.65 ± 0.01	−8.59 ± 0.02	−7.91 ± 0.03
14	t_{max} (MJD ± 0.5)	56176.666 ± 0.02	55846.320 ± 0.01	55928.414 ± 0.02
15	$t_{50\%,\text{after}}$ (days)	12.46 ± 0.01	14.08 ± 0.07	10.26 ± 0.01
16	Red-giant companion, percentage of angles excluded	100, 100, 100	100, 100, 100	72, 66, 61
17	Percentage of angles excluded for a six-solar-mass companion	94, 90, 86	100, 100, 98	0, 0, 0
18	Percentage of angles excluded for a two-solar-mass companion	77, 68, 61	94, 89, 84	0, 0, 0

The following properties are listed in the rows: (1) identifier; (2) number in the Kepler Input Catalog²⁶; (3) and (4) sky coordinates; (5) the Kepler magnitude of the galaxy; (6) galaxy redshift; (7) V-band Galactic dust extinction; (8) peak Kepler magnitude of the supernova; (9) absolute magnitude of the supernova, corrected for extinction and assuming typical cosmological parameters: $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.73$, $\Omega_{\text{matter}} = 0.27$; (10) Δ in the MLCS2k2 fit; (11), (12) and (13) the slope, exponent, and time of first detected light in the power-law fit, and the lower and upper 1σ bounds; (14), (15) and (16) the time of 50% light level before maximum, time of maximum, and the time of 50% level after maximum (the systematic error for t_{max} is about half a day); (17) the percentage of excluded viewing angles for red-giant companions (at the 68%, 95% and 99.7% confidence levels); (18) and (19) as in (17), but for a six-solar-mass and a two-solar-mass main-sequence companion. All listed times are with respect to maximum light, in rest-frame

¹Astronomy Department, University of Maryland, College Park, Maryland 20742-2421, USA. ²Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. ³Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA. ⁴Mt Stromlo Observatory, The Australian National University, via Cotter Road, Weston Creek, Australian Capital Territory 2611, Australia. ⁵Department of Astronomy, University of California, Berkeley, California 94720-3411, USA. ⁶Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA. ⁷Gemini Observatory, Southern Operations Center, c/o AURA, Casilla 603, La Serena, Chile.

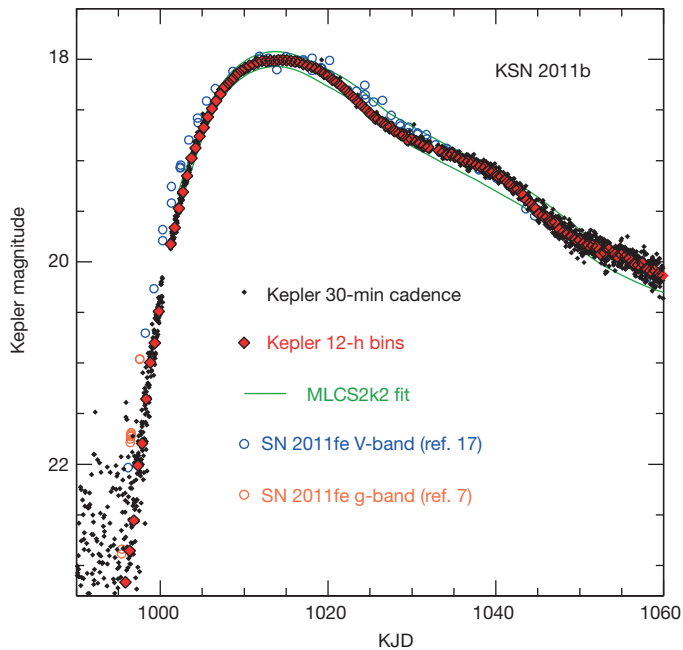


Figure 1 | Ground-based and Kepler lightcurves compared. The lightcurve of KSN 2011b was obtained by Kepler with a 30-min cadence (black) and binned at 12 h (red). The green lines indicate the range of best-fit type Ia supernova lightcurve templates (using the MLCS2k2¹⁹ analysis, modified to apply the Kepler wavelength sensitivity function). A secondary peak in the lightcurve, characteristic of type Ia supernovae, is apparent in KSN 2011b. Open circles show the lightcurve of the well-observed, nearby SN 2011fe^{7,17}, which has been shifted in magnitude and time to match the peak of KSN 2011b. SN 2011fe had a lightcurve width near the average for type Ia supernovae, whereas the KSN 2011b lightcurve is narrower. KJD measures time, in days, since 1 January of the year that the Kepler mission was launched (2009), with $KJD = MJD - 54832.5$.

We discovered the supernovae in the Kepler data by searching for variability resembling supernova lightcurves. Spectra of the host galaxies obtained subsequently with the Gemini and Keck telescopes indicate that all of our supernovae occurred in red passive galaxies with redshifts of about 0.1 (see Table 1).

The high quality of the Kepler data (Fig. 1) can be seen by comparing the lightcurves of Kepler supernova KSN 2011b and SN 2011fe^{7,17}, which has one of the best ground-based early supernova lightcurves. The Kepler photometry of KSN 2011b is comparable to that of SN 2011fe in depth, but with a significantly better cadence and overall photometric stability.

We compared the Kepler transients to supernova lightcurve templates using the PSNID code¹⁸, finding that KSN 2011b and KSN 2012a are clearly type Ia supernovae. The classification of KSN 2011c is more uncertain, with a PSNID 54% probability of being type Ia and a 46% probability of being type Ib (a core-collapse supernova). However, since the massive-star progenitor of a type Ib supernova is very unlikely to occur in an elliptical host galaxy, we classify KSN 2011c as a faint type Ia supernova. The three supernova lightcurves are also well fitted around their peaks by the type Ia supernova fitting program MLCS2k2¹⁹ (Fig. 2), supporting the type Ia supernova identification.

KSN 2012a and KSN 2011b both show a secondary ‘bump’ in the post-maximum lightcurve, a characteristic feature of normal type Ia supernovae observed at red wavelengths. KSN 2011c has the features of underluminous events: it lacks a clear second bump, rapidly declines, and is underluminous. KSN 2012a is moderately underluminous, while KSN 2011b is close to normal. Our type Ia supernovae are systematically offset from the brightness distribution typically observed for type Ia supernovae²⁰ (Fig. 2e); however, our Kepler galaxy selection was biased towards red, passive galaxies, which preferentially host dimmer supernovae²¹.

Simple analytic radiative-transfer models^{1,22} with power-law ejecta profiles predict that the early luminosity L can be described by a power law in time t as follows: $L \propto t^\alpha$, where $\alpha = 1.5-2$, depending on the

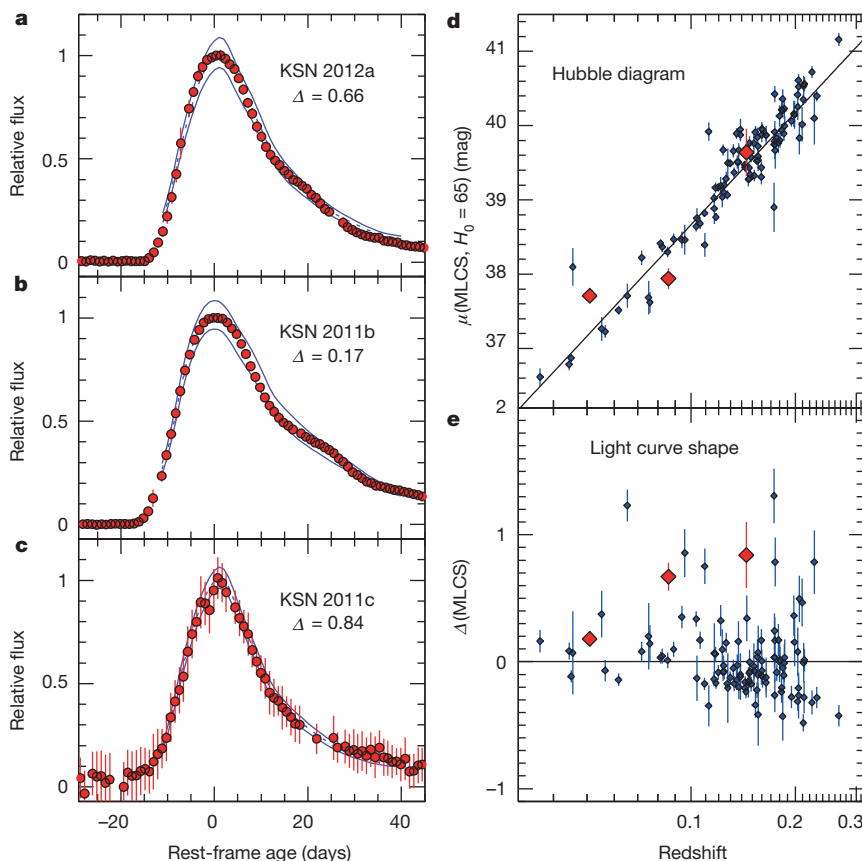


Figure 2 | Lightcurves of the three Kepler type Ia supernovae. The blue lines in a–c show the range of best-fit MLCS2k2¹⁹ lightcurve templates. All lightcurves are rather flat around maximum brightness, where they change little from about 1.5 days before to 1.5 days after maximum. **d**, Hubble diagram of the Kepler supernovae (red) along with type Ia supernovae from the SDSS-II survey²⁰ (blue), with μ the distance modulus in magnitudes. **e**, The distribution of the width parameter Δ versus redshift. The parameter Δ measures lightcurve shape and is anticorrelated with peak luminosity. The Kepler supernovae are biased towards large Δ because the Kepler galaxies have lower star-formation rates, which preferentially host dimmer supernovae²¹. All error bars are 1σ errors.

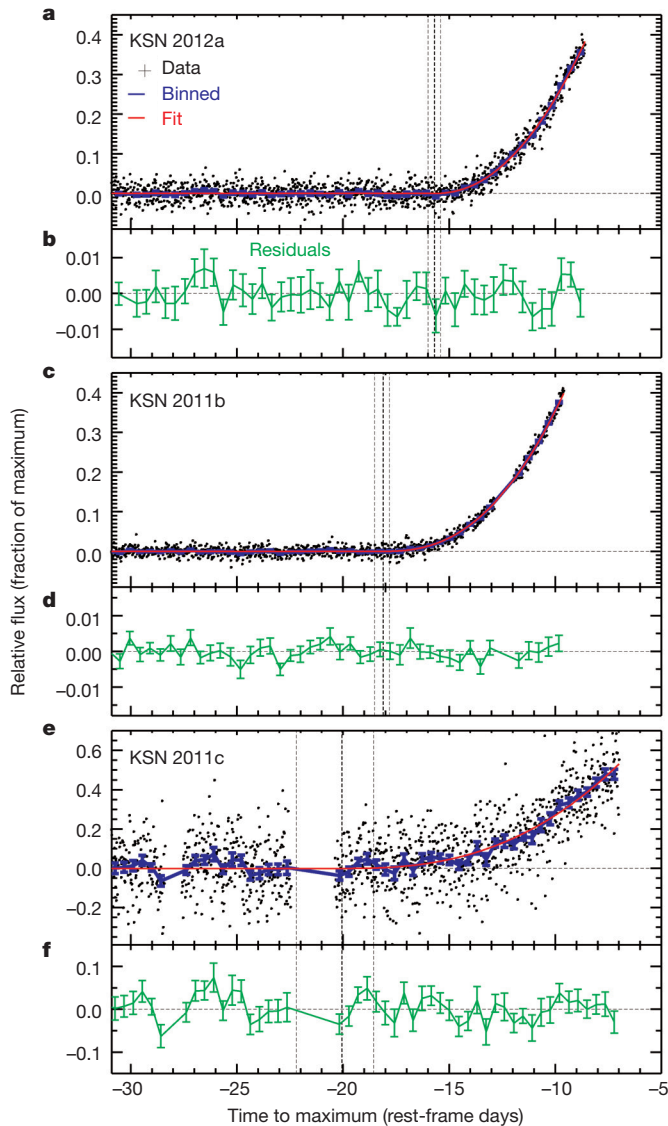


Figure 3 | The rise of the lightcurves. **a**, **c**, and **e** show the raw data (+ symbols), along with the data binned to half-day resolution (blue points and errors) and the best power-law fit (red line). The data are well fitted by a power law up to the time that the lightcurve reaches 40% of its peak value, or 7.2 days, 8.5 days, and 11.7 days for KSN 2012a, KSN 2011b, and KSN 2011c, respectively. The binned residuals (data minus fit) are shown in **b**, **d**, and **f** as green points with error bars. The time of first light and the $\pm 1\sigma$ uncertainty are shown as the vertical lines. In units of the peak supernova brightness, the unbinned errors are about 0.022, 0.012 and 0.120 for KSN 2012a, KSN 2011b and KSN 2011c, respectively. All error bars are 1σ errors.

ejecta structure. In detailed numerical models that use more complex ejecta structures, the early lightcurve need not follow a single power law. Previous studies^{11,23} of sparsely sampled lightcurves of numerous observed type Ia supernovae found an early rise consistent with a power law with $\alpha \approx 2$. However, the early lightcurves of SN 2013dy⁹ and SN 2014j¹⁰ are not well fitted by a single power law. Indeed, theoretical models predict that the shape of early type Ia supernova lightcurves is sensitive to the density and velocity structure of the supernova ejecta as well as the radial distribution of radioactive ^{56}Ni (ref. 1).

Figure 3 presents the Kepler early lightcurves, along with the best-fit power-law function $L(t) = C(t - t_0)^\alpha$, where t_0 is the time of first light and C is a constant, and where L is normalized by the peak supernova luminosity. We find that the data are all well fitted by a single power

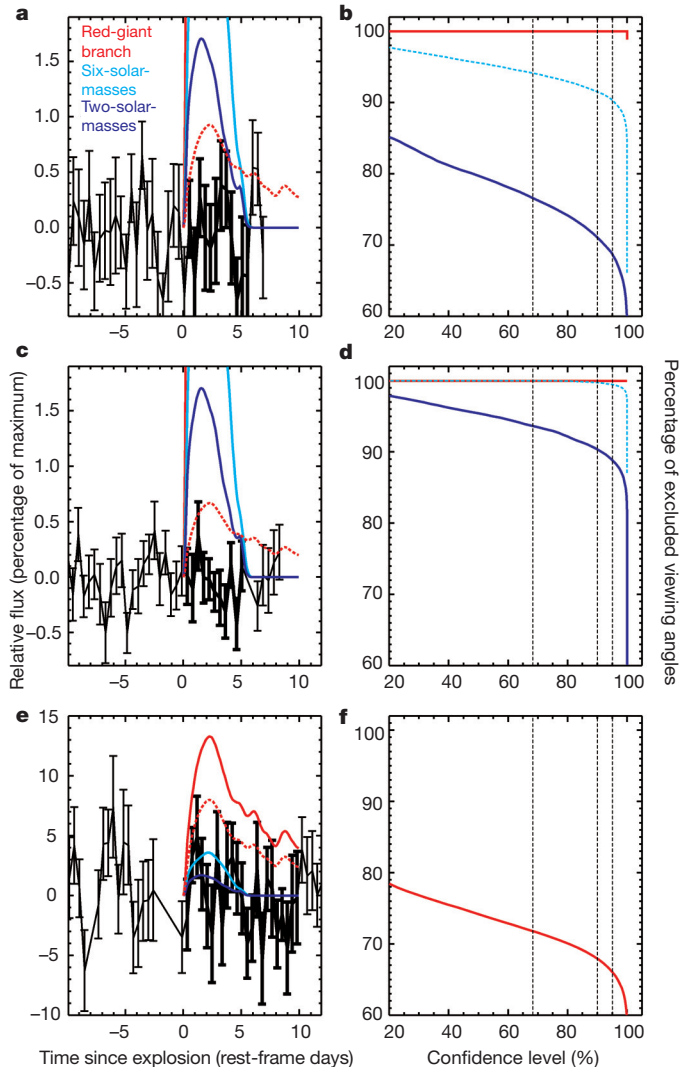


Figure 4 | Predicted maximum photometric signatures due to companions. **a–c** display the residuals of the power-law fit to the lightcurves (black points) and the average amount of shock emission predicted when the companion is either a red giant (red lines), or a six-solar-mass star (cyan) or a two-solar-mass main-sequence star (blue). The dashed red line shows the brightest possible red-giant shock contribution at the 68% confidence level. We simulate observing from different viewing angles by scaling the shock luminosity, and determining the probability that such a model can be excluded. **d–f** show the percentage of viewing angles that can be excluded at a given confidence level. All error bars are 1σ errors.

law up to the point that $L(t)$ reaches 40% of the peak, beyond which the lightcurve turns over. The indices are $\alpha = 2.12 \pm 0.14$, 2.44 ± 0.14 , and 2.58 ± 0.33 for KSN 2012a, KSN 2011b and KSN 2011c, respectively. The errors include the correlations between the fitted parameters (see Methods and Extended Data Fig. 1). The weighted average index of 2.3 ± 0.09 differs from the value $\alpha \approx 2$ found in some previous studies of supernovae observed at much lower cadence^{11,23} (but see SN 2013dy⁹ and SN 2014j¹⁰ for additional diversity in early behaviour). The Kepler supernovae thus provide new constraints on theoretical models that link the lightcurve rise to the properties of the supernova ejecta and the explosion mechanism.

The shape of the early lightcurve can also be affected by emission resulting from the collision of the supernova ejecta with circumstellar matter or a companion star. Theoretical models² find that the excess optical/ultraviolet emission due to the collision shock causes a deviation from a simple power-law rise. The luminosity of the shock emission is directly proportional to the radius of the companion star

(assumed to be in Roche-lobe overflow) and is dependent on orientation, being brightest when the observer's viewing angle is aligned with the region of shocked ejecta. The lightcurves of SN 2011fe and SN 2013dy show no signs of a shock interaction. Pre-explosion images²⁴ of SN 2011fe taken with the Hubble Space Telescope exclude a red-giant or subgiant²⁵ companion larger than about four solar masses.

In Fig. 4a, c and e, we examine the residuals of the power-law fits of the Kepler supernovae and find no systematic trends indicative of shock emission. For comparison, we overplot predictions² of the shock emission for three models assuming companions with orbital separations of 2×10^{13} cm (typical for a red giant), 2×10^{12} cm (typical of a main-sequence star of six solar masses) and 5×10^{11} cm (a two-solar-mass star). The residuals of KSN 2012a and KSN 2011b are inconsistent with a red-giant companion viewed from any viewing angle. There is also no indication of a main-sequence star for most viewing angles. The fractions of viewing angles excluded for the six-solar-mass and two-solar-mass models are, respectively, 94% and 77% for KSN 2012a and 100% and 94% for KSN 2011b, at the 68% confidence level. KSN 2011c is less constraining, given the lower signal-to-noise ratio.

Our discovery of three type Ia supernovae with Kepler has opened a new window on the progenitor system and explosion physics. K2, the follow up to the Kepler mission, is primarily used to search for exoplanets. We use K2 to monitor up to several thousand galaxies, which are being simultaneously observed by ground-based photometric and spectroscopic programmes, making it highly likely that many more supernovae will be detected in the early stages of the explosion and allowing early and detailed follow-up observations.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 January; accepted 23 March 2015.

- Piro, A. L. & Nakar, E. What can we learn from the rising light curves of radioactively powered supernovae. *Astrophys. J.* **769**, 67–79 (2013).
- Kasen, D. Seeing the collision of a supernova with its companion star. *Astrophys. J.* **708**, 1025–1031 (2010).
- Haas, M. *et al.* Kepler science operations. *Astrophys. J.* **713**, L115–L119 (2010).
- Foley, R. J. *et al.* Very early ultraviolet and optical observations of the SN Ia supernova 2009ig. *Astrophys. J.* **744**, 38 (2012).
- Hachinger, S. *et al.* The UV/optical spectra of the SN Ia supernova SN 2010jn: a bright supernova with outer layers rich in iron-group elements. *Mon. Not. R. Astron. Soc.* **429**, 2228–2248 (2013).
- Nugent, P. *et al.* Young SN Ia supernova PTF11kly in M101. *Astron. Telegr.* **3581**, 1 (2011).
- Nugent, P. E. *et al.* Supernova SN 2011fe from an exploding carbon-oxygen white dwarf star. *Nature* **480**, 344–347 (2011).
- Silverman, J. M. *et al.* The very young SN Ia supernova 2012cg: discovery and early-time follow-up observations. *Astrophys. J.* **756**, L7–L12 (2012).
- Zheng, W. *et al.* The very young SN Ia supernova 2013dy: discovery, and strong carbon absorption in early-time spectra. *Astrophys. J.* **778**, L15–L20 (2013).
- Zheng, W. *et al.* Estimating the first-light time of the type Ia supernova 2014J in M82. *Astrophys. J.* **783**, L24–L28 (2014).
- Hayden, B. T. *et al.* The rise and fall of SN Ia supernova light curves in the SDSS-II supernova survey. *Astrophys. J.* **712**, 350–366 (2010).
- Hayden, B. T. *et al.* Single or double degenerate progenitors? searching for shock emission in the SDSS-II type Ia supernovae. *Astrophys. J.* **722**, 1691–1698 (2010).
- Bianco, F. B. *et al.* Constraining SN Ia supernovae progenitors from three years of supernova legacy survey data. *Astrophys. J.* **741**, 20–31 (2011).
- Ganeshalingam, M., Li, W. & Filippenko, A. V. The rise-time distribution of nearby type Ia supernovae. *Mon. Not. R. Astron. Soc.* **416**, 2607–2622 (2011).
- Tucker, B. E. UV properties of type Ia supernova and their host galaxies. *Astrophys. Space Sci.* **335**, 223–230 (2011).
- Jenkins, J. J. *et al.* Initial characteristics of Kepler long cadence data for detecting transiting planets. *Astrophys. J.* **713**, L120–L125 (2010).
- Munari, U. *et al.* BVRI lightcurves of supernovae SN 2011fe in M101, SN 2012aw in M95, and SN 2012cg in NGC 4424. *New Astron.* **20**, 30–37 (2013).
- Sako, M. *et al.* Photometric type Ia supernova candidates from the three-year SDSS-II SN survey data. *Astrophys. J.* **738**, 162–177 (2011).
- Jha, S., Riess, A. G. & Kirshner, R. P. Improved distances to type Ia supernovae with multicolor light-curve shapes: MLCS2k2. *Astrophys. J.* **659**, 122 (2007).
- Kessler, R. *et al.* First-year Sloan Digital Sky Survey-II supernova results: Hubble diagram and cosmological parameters. *Astrophys. J.* **185** (Suppl.), 32–84 (2009).
- Gallagher, J. S. *et al.* Supernovae in early-type galaxies: directly connecting age and metallicity with type Ia luminosity. *Astrophys. J.* **685**, 752 (2008).
- Arnett, W. D. Type I supernovae. I.—Analytic solutions for the early part of the light curve. *Astrophys. J.* **253**, 785–797 (1982).
- Conley, A. *et al.* The rise time of type Ia supernovae from the Supernova Legacy Survey. *Astron. J.* **132**, 1707–1713 (2006).
- Li, W. *et al.* Exclusion of a luminous red giant as a companion star to the progenitor of supernova SN 2011fe. *Nature* **480**, 348–350 (2011).
- Bloom, J. S. *et al.* A compact degenerate primary-star progenitor of SN 2011fe. *Astrophys. J.* **744**, L17–L21 (2012).
- Brown, T. *et al.* Kepler Input Catalog: photometric calibration and stellar classification. *Astron. J.* **142**, 112–130 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Still, M. Fanelli, and S. Gezari for useful conversations, M. Graham, P. Kelly, K. Clubb, and O. Fox for assistance with the observations and reductions of the host-galaxy spectra, D. Scolnic for help with PSNID, and F. Bianco for sending us time series of the companion shock models integrated over the Kepler bandpass. D. Thilker kindly provided the SDSS magnitudes for the supernova host galaxies. R.P.O. and E.J.S. were, in a small part, supported by Kepler G03 and G04 grants NNX12AC95G and NNX13AC27G. P.M.G. was partly supported by Kepler grants NNX12AC89G and NNX11AG95G. A.V.F. and B.E.T. were supported by NSF grant AST-1211916, the TABASGO Foundation, and the Christopher R. Redlich Fund. Some of the data presented herein were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts. This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. This work is based in part on observations obtained at the Gemini Observatory (program IDs GN-2013A-Q-4 and GS-2013A-Q-115) which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil) and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina). Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA; the Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

Author Contributions R.P.O., R.M., and E.J.S. developed the idea of looking for variability in galaxies using Kepler data. R.P.O. selected the target galaxies and created the pipeline to analyse the Kepler data, and developed and implemented most of the innovations in data reduction. The analysis by R.P.O. and E.J.S. of Kepler's Full Frame Images (not discussed here) convinced us that long-term stability for Kepler data was achievable. E.J.S. has also confirmed the reduction results using independent techniques. P.M.G. modified the MLCS2k2 program to fit Kepler light curves. A.R. used PSNID to classify the supernova. D.K. computed the companion shock models. B.E.T. coordinated, with S.M. and A.V.F., the spectroscopic observations of the host galaxies and measured the redshifts. All authors contributed to the analysis and interpretation of the Kepler supernova lightcurves, as well as the text of this Letter.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.P.O. (olling@astro.umd.edu).

METHODS

Photometric calibration. The Kepler Mission^{3,16} provided nearly continuous observations of many galaxies, resulting in a few supernova lightcurves that span from months before to months after the supernova event at a 30-min cadence. Our Kepler guest observer projects, GO20058, GO30032, and GO40057, monitored about 400 galaxies at 30-min cadence to look for brightness variations in their nuclei indicative of an active galactic nucleus and to search for supernovae. Targets were selected from the 2MASS extended source catalogue²⁷. Typically, we obtained 8×8 pixel cutout regions centred on each galaxy for two to three years.

On a timescale of minutes to hours, Kepler provides photometric precision of a few parts in a million for bright sources^{3,16}. On longer timescales, various systematic effects considerably reduce the precision of the standard Kepler products. We have developed a specialized analysis which obtains uncertainties of a few parts in ten thousand on timescales from hours to years, at 17th magnitude.

The Kepler Archive Manual²⁸ describes where the data are stored and how the Presearch Data Conditioning (PDC) Maximum A Posteriori (MAP) lightcurve production procedure significantly reduces sensitivity variations on timescales exceeding a few days. However, PDC-MAP fails for objects having large intrinsic, astrophysical variations like supernovae. Below we describe our data-reduction procedures that reduce instrumental variations on long timescales while preserving real astrophysical signals. Our data-reduction method is based on the small cutout images that Kepler obtains. These so-called target (TARG) pixel images have undergone standard astronomical calibration, including background subtraction and flagging of cosmic-ray events or otherwise bad data²⁸. The TARG files contain the background-subtracted images, as well as the subtracted background itself (FLUX_BKG). Kepler's simple aperture photometry (SAP) lightcurves are extracted from these TARG images where only the light in certain pixels is added together. For the SAP, the choice of pixels, chosen to optimize the signal-to-noise ratio at short timescales, is based on the source brightness and the shape of the point-spread function.

The Kepler data processing is organized in three-month chunks labelled quarters Q0 to Q17. About once per month, the spacecraft goes through a pointing manoeuvre to downlink the data to Earth. We virtually eliminate the large spikes in the SAP lightcurves after repointing manoeuvres by using larger apertures. We use three different apertures to measure the lightcurves. The first aperture type ('5×5') uses a 5×5 pixel box centred on our galaxies. For the second type of aperture ('WEI'), we identify which pixels contain most of the light from the galaxy but are not contaminated badly by neighbouring sources, and we assign a high weight to pixels near the source centre and lower weights to pixels further out, decreasing to zero near strong confusing sources. Pixels without contributions from sources are given intermediate weight but are separately identified as background pixels. We then create a curve of growth with the cumulative flux as a function of pixel weight. If an error has been made in determining the background level, the upper part of the curve of growth will tilt upward or downward. Confusing sources will add their own growth curve. This procedure generally produces nonvarying lightcurves for most of our galaxies, as expected. The WEI apertures typically contain 30 to 40 pixels.

Our third aperture type ('WAL') uses the ensemble of the WEI apertures from all quarters to design an aperture with the same number of pixels in all quarters. Lightcurves produced with these apertures also yield flat lightcurves for the majority of our galaxies. Close inspection of results for our ~ 400 galaxies indicates that it is not possible to decide which of the 5×5, WEI, or WAL lightcurves is best, so we decided to combine all three apertures and use the variance amongst them as an indicator of the systematic uncertainties.

Another issue is that the lightcurves do not match up at the boundaries of quarters. A new quarter brings a new orientation of the Kepler spacecraft and thus a different charge-coupled device (CCD). This has two effects: (1) even for well-calibrated CCDs, there are sensitivity variations of a few per cent between pixels, and (2) an aperture in the next quarter could have a different size and be centred differently with respect to the source centre. Consequently, the apertures capture a different fraction of the total light of the source. Both effects indicate that the quarters need to be 'stitched together' by using a multiplicative factor. Specifically, we fitted a polynomial to the last few days of the previous quarter and the first few days of the next quarter, thereby determining a multiplicative factor. Our objects could be quiet or rapidly changing at a quarter boundary; therefore, we used time windows of 4–22 days and polynomial fits up to fifth order, and used the combination of parameters that had the smallest residuals. This procedure works well but is imperfect, perhaps because of background-subtraction errors; these could be removed in future analyses with an additive stitching factor.

The remaining photometric trends strongly correlate with the background levels as extracted from the 'FLUX_BKG' in the target pixel files. This suggests that the background levels are not optimally determined by the Kepler pipeline. Based on the weights determined for the 'WEI' aperture above, we identify a

number of pixels in each cutout as 'background pixels', and we use those pixels to construct our own estimate of the background level. We think that these 'backgrounds' are most probably caused by zodiacal emission from dust in the Solar System, and not by faint galaxies or stars. We used the model from the COBE/DIRBE team to predict the zodiacal emission²⁹ at a wavelength of $1.25 \mu\text{m}$ towards the Kepler field, as seen from Kepler. There is a remarkable correspondence between the scaled COBE/DIRBE model and the background levels from the Kepler data, justifying the removal of structure in our lightcurves that follows the zodiacal emission. However, the long-term variation in our lightcurves does not exactly follow the zodiacal emission. Finally, the zodiacal emission may exhibit variations that do not follow the model, possibly as a result of interactions between the zodiacal dust and coronal mass ejections, the solar wind, or the intersection of old comet dust trails with Kepler's line of sight. In fact, brightening of the background lasting days to weeks has been observed.

The trends in counts of our ~ 400 galaxies vary in roughly the same fashion, exhibiting four significant sinusoidal components with almost identical periods and phases. The periods of these components are about one, half, a third, and a quarter of a Kepler year. Coefficients for these four sinusoids were determined for the quiet periods of each supernova lightcurve and the sinusoids were subtracted from the entire time series.

We use three different versions of the lightcurves and fit two different versions of our four-component sinusoidal model to the long-term background variations. For each of our 5×5, WEI, and WAL apertures we thus have six ways to eliminate the background effects. This leads to 18 different estimates of the actual, intrinsic lightcurve. As far as we can determine, none of these versions is superior to the others, so we construct a median value and its root-mean-square (RMS) 'error' of the 18 different estimates of each observation, after clipping outliers. When we present binned data in the figures or the data tables, we use the median value of typically 26 data points (corresponding to one-half day in the observed frame) where we also reject outliers and undefined values. The reported errors are the weighted RMS scatter of the data, divided by the square-root of the number of valid data points.

The three different versions of the lightcurves are (1) the raw lightcurve being either 5×5, WEI, or WAL; (2) the raw lightcurve minus a scaled version of the Kepler Project's background, where the scale factor is chosen so as to minimize the residuals; and (3) the raw lightcurve minus a scaled version of our own background. The two different fits we perform on raw lightcurves 1–3 are: (1) a four-component sinusoidal fit where the periods are set to the one, half, a third and a quarter-year periods but allowing for a few days variation in those periods, and (2) as above, but allowing the periods to change by $\pm 7\%$. In these fits, we also include a quadratic polynomial to accommodate the longest trends, including sensitivity losses.

Inspection of a large number of galaxy lightcurves produced in this way indicates that the procedure works well: long-term trends are removed and residuals are around the 0.25% level. In some cases, however, larger errors still occur, mostly at the boundaries between quarters. The lightcurves for the supernovae presented here extend well beyond the regions that we used in the main part of this Letter, and we do not see signs of problems there. Thus, we are confident that the lightcurves presented here are free from systematic effects above the 0.25% level on timescales of weeks to years. As a function of Kepler magnitude (K_p), we achieve errors of $3.6 + 0.29(K_p - 18)^4$ millimagnitudes, on a timescale of half a day. However, small residual background variations may still be present on timescales of days to weeks, and we take this into account via the background B term in equation (2) below.

Lightcurve fitting. We parameterize the observed lightcurve L , normalized by its maximum, as the sum of the supernova part (L_{SN}) and a small background term, $L(t) = L_{\text{SN}}(t) + B(t)$, with

$$L_{\text{SN}}(t) = C(t - t_0)^\alpha \quad (1)$$

and

$$B(t) = b_0 + b_1(t - t_0) + b_2(t - t_0)^2 \quad (2)$$

The background terms are fitted over the time range before the onset of the event, and the $L(t)$ function is fitted where $t \geq t_0$, but with the parameters of the background fixed. The time range used for fitting starts roughly 20 days (30 days for KSN 2011c) before the supernova explosion and lasts until the lightcurve reaches 40% of the peak amplitude, that is, until $t = t_{(L_{\text{SN}} = 0.4)}$. The maximum occurs at $t = t_{\text{max}}$. These ranges have been empirically determined to ensure that (1) $B(t)$ faithfully describes the region before the supernova and is not affected by any variations lasting less than about one week; and (2) the eventual turnover of the lightcurve does not affect the fitted power-law index. The fits to the early

lightcurves are excellent, and the reduced- χ^2 values confirm that this model is a good match to the data.

Experimentation with synthetic lightcurves that represent well the overall supernova lightcurve yield constant α values, as long as the time span used for fitting is less than about $t_{\text{LSN}} = 0.4$. When using longer time spans, the fitted α and t_0 change systematically. This effect may explain why different authors have found different α values^{9–11,23}: different time spans yield different power-law indices.

The reported uncertainties on t_0 , α , and C are derived from $N_{\text{mod}} = 2 \times 10^6$ random realizations of the data, where we perturb each measurement by a random value based on the photometric error of the measurement, and where we perturb the initial estimates for the fitting function by the a posteriori 1σ errors. To incorporate the strong correlations between the fit parameters, we determine the parameter uncertainties from the maximum extent of the error ellipses that contain 68% of all models. The error ellipses are determined from the ensemble of the two-dimensional distribution (contour plots) of all the fit results (Extended Data Fig. 1). These maximal extents of the 68%-error ellipses are indicated by the horizontal and vertical lines in Extended Data Fig. 1.

Each contour diagram in Extended Data Fig. 1 comprising parameters i and j is a measure of the total number of fitted models that fall within pixel (k, l) with parameter values (v_i, v_j) . The contours drawn are based on the cumulative distribution of the pixel values, starting from the largest value. For example, the 0% contour corresponds to the peak in the map, the 100% contour encloses all pixels that contain fit results, while the 25% contour contains all pixels whose sum equals a quarter of the number of models. Each map then provides an estimate for the average values A and the lower (L) and upper (U) 1σ error bounds. We compute the average of parameter i based on map (i, j) as:

$$A_{i,j} = \{ \sum_{k,l} [v_i \times N_{k,l}] / \sum_{k,l} [N_{k,l}] \} \quad (3)$$

and similarly for $A_{j,i}$, where the sum over k, l includes all pixels that contain 25% of all fit results. The lower 1σ bound of parameter i , $L_{i,j}$, corresponds to the lower projection of the 68% contour onto the i axis, and likewise for the upper bound $U_{i,j}$. The symmetric error, $e_{i,j}$, equals $(U_{i,j} - L_{i,j})/2$. Since there are three fit parameters, each parameter i is determined from two contour maps, (i, n) and (i, m) . The weighted average for parameter i is then:

$$\langle A_i \rangle = \left(\sum_{j=n,m} w_{i,j} \times A_{i,j} \right) / \left(\sum_{j=n,m} w_{i,j} \right) \quad (4)$$

where the weight $w_{i,j} = (1/e_{i,j})^2$ and the weighted error on $\langle A_i \rangle$ is:

$$e_i = \left(\sum_{j=n,m} w_{i,j} \right)^{-1/2} \quad (5)$$

These lower, average, and upper values are determined for 5,000 subsets of the total number of models, and averaged.

Extended Data Fig. 1 shows the confidence regions for the α - t_0 and α - C correlations, for all two million fit results combined (the t_0 - C map looks similar). All parameters are highly correlated, which leads to relatively large and somewhat asymmetric errors. The results are given in Table 1.

We have performed extensive simulations of this fitting process, on both synthetic and actual data, to arrive at this procedure. Details will be presented elsewhere; here we discuss a few highlights. Because the curvature of the χ^2 surface is low and riddled with local minima, our fitter (modified Levenberg–Marquardt) can easily get stuck in a local minimum. Consequently, the fitted values depend on the initial guess for the fit parameters—that is, if the initial estimates are off, the results will be biased. To avoid these problems, we determine robust initial estimates, and randomize the initial guess values by an amount that roughly corresponds to the a posteriori errors. The fit results for KSN 2011c may have additional systematic errors of unknown magnitude owing to our quarter-stitching procedure.

Figure 3 and the fitted values listed in Table 1 indicate that the early lightcurves of the Kepler supernovae are substantially different: the power-law indices differ by 21%, the times to reach 40% of the maximum vary by 55%, and the C terms vary by a factor of 8.4. Thus, there are substantial different evolutions of the temperature, the opacity, the rate of expansion of the ejecta, and the internal distribution of radioactive elements¹.

Shockwave interaction with the companion. The absolute strength of the shock emission depends on the size of the companion and its distance from the exploding white dwarf². Analysis of the model light curves indicates that the observable shock emission depends on two orthogonal parameters: (1) the normalized time dependence of the shock emission, $I_{\text{SE}}(t)$, which is rather similar between 0.5 day and 5.5 days for models that contain either a red giant or a main-sequence companion (Fig. 4); and (2) the viewing angle θ away from the shocked region. The observable

shock emission is then reasonably well approximated by $\gamma_{\text{obs}}(t, \theta) = I_{\text{SE}}(t) \times S(\theta)$, where, to within the errors:

$$S(\theta) \approx 0.982 \times \exp(-(\theta/99.7)^\alpha) + 0.018 \quad (6)$$

with θ in degrees, and where the minimum observed shock strength of 0.055 occurs at $\theta = 180^\circ$.

To evaluate the maximum strength allowed by the data, we multiply the model $I_{\text{SE}}(t)$ curves by a series of factors s and compute the resulting χ^2_s value and probability P_s that the data are consistent with $\gamma_{\text{obs}}(t, \theta)$. As an example, in Fig. 4a, c and e we plot (red dashed line) the strongest shock model that is consistent with the data, at 68% confidence.

Thus, each s value corresponds to a level of confidence $C_s = (1 - P_s)$ that a model can be excluded by the data. Likewise, each s value also corresponds to a viewing angle θ (through equation (6)). Finally, given the three-dimensional distribution of viewing angles, the fraction of viewing angles that have strength exceeding $s = S(\theta)$ is $f(S) = (1 - \cos\theta)/2$. In Fig. 4b, d and e, we present $f(S)$ as a function of C_s ; the fraction of excluded viewing angles (too strong shocks) as a function of the confidence level that a shock with that strength can be excluded by the data.

Lightcurve analysis. We analysed the lightcurves of the Kepler supernovae using the fitting program MLCS2k2 (ref. 19). In MLCS2k2 a set of templates has been created from a large number of real type Ia supernova light curves, and K-corrections are calculated from a large sample of observed spectra. A Kepler bandpass was converted to the MLCS2k2 format and used to K-correct the Kepler magnitudes to the standard rest-frame R band. MLCS2k2 also fits the colour curves to estimate reddening due to dust in the host galaxy. However, there is no colour information for the Kepler supernovae, so the extinction was fixed at zero.

Kepler magnitudes of the host galaxies. Both the 'SDSS' and Kepler magnitudes listed in the Kepler Input Catalog²⁶ for these supernova host galaxies were based on the photographic magnitudes from the USNO-B Catalog³¹, which were transformed to SDSS-like magnitudes for the Kepler Input Catalog²⁶. Such photographic magnitudes rely on the diameter of the 'spots' on the photographic plates, which are affected by the fact that these sources are extended. Accordingly, the Kepler magnitudes listed in the Kepler Input Catalog²⁶ for our galaxies are much too bright. The Kepler magnitudes listed in Table 1 are computed using CCD-based SDSS magnitudes.

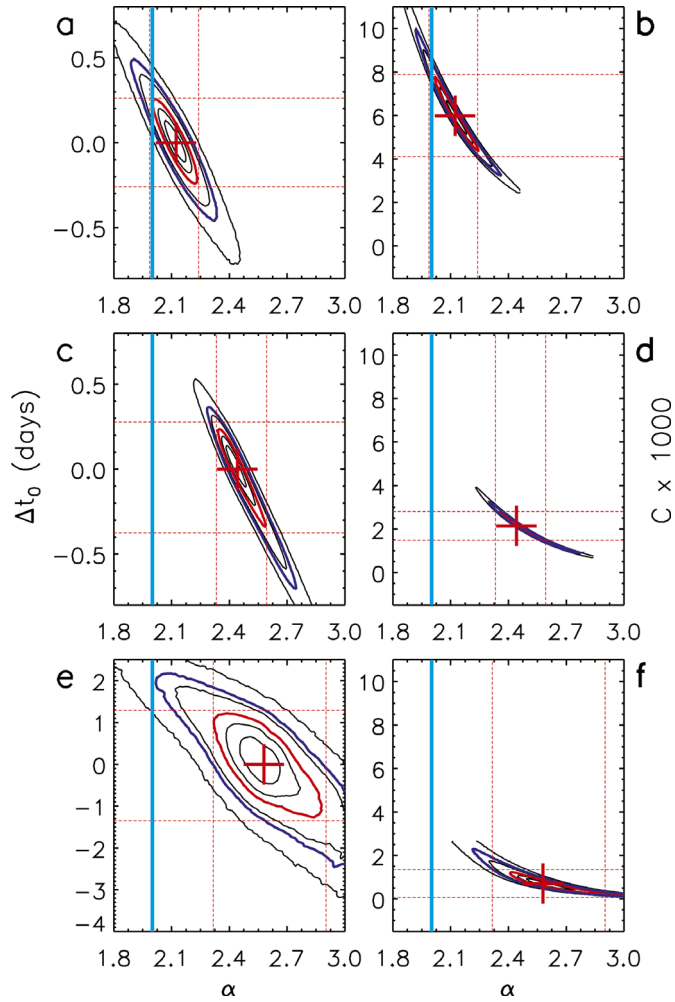
Spectroscopy of the host galaxies. Spectra were taken of the host galaxies of our three type Ia supernovae. Spectra obtained at Gemini North used the Gemini Multi-Object Spectrograph, with the R400 grating and the 1.5" slit. At Keck, we used the Low Resolution Imaging Spectrometer³⁰ with the 600/4000 grism and 400/8500 grating for coverage of wavelengths 3,010–9,000 Å. All data were reduced using standard IRAF³² tasks to remove instrument and sky signatures, and to extract one-dimensional spectra for analysis.

The host galaxy of KSN 2011b is a red, passive galaxy at redshift $z = 0.052$. It has a high [N II]/H α ratio, indicative of a possibly active nucleus, and we detect [S II] emission and Ca II H and K absorption. KSN 2011c occurred in a red, passive galaxy at $z = 0.144$; we see Ca II H and K, Mg II, Na I, and multiple Fe I and Fe II absorption lines. KSN 2012a was also in a red, passive galaxy, at $z = 0.086$. The spectrum shows multiple absorption features (Ca II H&K, Mg II, Na I), and the 4,000 Å break.

Code availability. Except for the programs mentioned below, most code is written in IDL by R.P.O., but is not available for public consumption due to its many undocumented intricacies, and the lack of funds to make it possible to publish the code in a user-friendly form. PSNID and MLCS2k2 can be obtained at: <http://www.sas.upenn.edu/~gladney/html-physics/psnid/psnidIII/> and <http://www.physics.rutgers.edu/~saurabh/mlcs2k2/>, respectively. The shock interaction models were taken from the previously published results of ref. 2, which were run with the radiation transport code SEDONA (<http://adsabs.harvard.edu/abs/2006ApJ...651..366K>), which is not currently available for open access.

Source data for figs 1–4 are provided in the Supplementary Information.

27. Jarrett, T. H. *et al.* 2MASS Extended Source Catalog: overview and algorithms. *Astron. J.* **119**, 2498–2531 (2000).
28. Thompson, S. E. *et al.* *Kepler Archive Manual* http://archive.stsci.edu/kepler/manuals/archive_manual.pdf (5 June 2014).
29. Kelsall, T. *et al.* The COBE diffuse infrared background experiment search for the cosmic infrared background. II. Model of the interplanetary dust cloud. *Astron. J.* **508**, 44–73 (1998).
30. Oke, J. B. *et al.* The Keck low-resolution imaging spectrometer. *Publ. Astron. Soc. Pacif.* **107**, 375–385 (1995).
31. Monet, D. *et al.* The USNO-B Catalog. *Astron. J.* **125**, 984–993 (2003).
32. Tody, D. *Astronomical Data Analysis Software and Systems II* (eds Hanisch, R. J., Brissenden, R. J. V. & Barnes, J.), Vol. 52 of *Astronomical Society of the Pacific Conference Series*, 173–183 (1993).



Extended Data Figure 1 | Confidence regions of fitted parameters. Two-dimensional projections (contour maps) of the three-dimensional distribution of the fit parameters α , t_0 , and C . For Kepler supernovae 2012a, 2011b, and 2011c, from top to bottom. **a**, **c**, and **e** show α versus t_0 , and **b**, **d**, and **f** display α versus C . The contours, from inside to out, contain 25%, 50%, 68.3% (red), 90%, 95.5% (blue), and 99.7% of all 2×10^6 Monte Carlo model-fit results. The dashed red lines are the $\pm 1\sigma$ limits of the projections. The vertical cyan line represents the fireball model²². Details of the fitting procedure are described in the Methods section.

Spin–motion entanglement and state diagnosis with squeezed oscillator wavepackets

Hsiang-Yu Lo¹, Daniel Kienzler¹, Ludwig de Clercq¹, Matteo Marinelli¹, Vlad Negnevitsky¹, Ben C. Keitch^{1†} & Jonathan P. Home¹

Mesoscopic superpositions of distinguishable coherent states provide an analogue of the ‘Schrödinger’s cat’ thought experiment^{1,2}. For mechanical oscillators these have primarily been realized using coherent wavepackets, for which the distinguishability arises as a result of the spatial separation of the superposed states^{3–5}. Here we demonstrate superpositions composed of squeezed wavepackets, which we generate by applying an internal-state-dependent force to a single trapped ion initialized in a squeezed vacuum state with nine decibel reduction in the quadrature variance. This allows us to characterize the initial squeezed wavepacket by monitoring the onset of spin–motion entanglement, and to verify the evolution of the number states of the oscillator as a function of the duration of the force. In both cases we observe clear differences between displacements aligned with the squeezed and anti-squeezed axes. We observe coherent revivals when inverting the state-dependent force after separating the wavepackets by more than 19 times the ground-state root mean squared extent, which corresponds to 56 times the root mean squared extent of the squeezed wavepacket along the displacement direction. Aside from their fundamental nature, these states may be useful for quantum metrology⁶ or quantum information processing with continuous variables^{7–9}.

The creation and study of non-classical states of spin systems coupled to a harmonic oscillator have provided fundamental insights into the nature of decoherence and the quantum–classical transition. These states and their control form the basis of experimental developments in quantum information processing and quantum metrology^{1,2,10}. Two of the most commonly considered states of the oscillator are squeezed states and superpositions of coherent states of opposite phase, which are commonly referred to as ‘Schrödinger’s cat’ (SC) states. Squeezed states involve a reduction of the fluctuations in one quadrature of the oscillator below the ground-state uncertainty, which has been used to increase sensitivity in interferometers^{11,12}. SC states provide a complementary sensitivity to environmental influences by separating the two parts of the state by a large distance in phase space. These states have been created in microwave and optical cavities^{2,13}, where they are typically not entangled with another system, and also with trapped ions^{1,3–5}, where all experiments performed have involved entanglement between the oscillator state and the internal electronic states of the ion. SC states have recently been used as sensitive detectors for photon scattering recoil events at the single-photon level¹⁴.

Here we use state-dependent forces (SDFs) to create superpositions of distinct squeezed oscillator wavepackets that are entangled with a pseudo-spin encoded in the electronic states of a single trapped ion. We will refer to these states as squeezed wavepacket entangled states (SWESs) in the rest of the paper. By monitoring the spin evolution as the entanglement with the oscillator increases^{15–17}, we are able to observe the squeezed nature of the initial state directly. We obtain a complementary measurement of the initial state by extracting the number-state probability distribution of the displaced-squeezed states that make up the superposition. In both measurements we observe

clear differences depending on the force direction. We show that the SWESs are coherent by reversing the effect of the SDF, resulting in recombination of the squeezed wavepackets, which we measure through the revival of the spin coherence.

The squeezed vacuum state $|\xi\rangle$ is defined by the action of the squeezing operator $\hat{S}(\xi) = e^{(\xi^* \hat{a}^2 - \xi \hat{a}^{\dagger 2})/2}$ on the motional ground state $|0\rangle$, where $\xi = re^{i\phi_s}$, with r and ϕ_s real parameters that define the magnitude and the direction of the squeezing in phase space. To prepare squeezed states of motion in which the variance of the squeezed quadrature is reduced by about 9 dB relative to the ground-state wavepacket we use reservoir engineering, in which a bichromatic light field is used to couple the ion’s motion to the spin states of the ion, which undergo continuous optical pumping. This dissipatively pumps the motional state of the ion into the desired squeezed state, which is the dark state of the dynamics. More details about the reservoir engineering can be found in ref. 18. This approach provides a robust basis for all experiments described below, typically requiring no recalibration over several hours of taking data. In the ideal case, the optical pumping used in the reservoir engineering results in the ion being pumped to $|\downarrow\rangle$. To create a SWES, we apply a SDF to this squeezed vacuum state by simultaneously driving the red $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n-1\rangle$ and blue $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n+1\rangle$ motional sidebands of the spin-flip transition³. The resulting interaction Hamiltonian can be written in the Lamb–Dicke approximation (LDA) as

$$\hat{H}_D = \hbar \frac{\Omega}{2} \hat{\sigma}_x (\hat{a}^\dagger e^{-i\phi_D/2} + \hat{a} e^{i\phi_D/2}) \quad (1)$$

where Ω is the strength of the SDF, ϕ_D is the relative phase of the two light fields, and $\hat{\sigma}_x \equiv |+\rangle\langle+| - |-\rangle\langle-|$ with $|\pm\rangle = (|\uparrow\rangle \pm |\downarrow\rangle)/\sqrt{2}$. For an ion prepared in $|+\rangle$, this Hamiltonian results in displacement of the motional state in phase space by an amount $\alpha(\tau) = -i\Omega e^{-i\phi_D/2} \tau/2$, which is given in units of the root mean squared (r.m.s.) extent of the harmonic oscillator ground state. An ion prepared in $|-\rangle$ will be displaced by the same amount in the opposite direction. In the following equations we use α in place of $\alpha(\tau)$ for simplicity. Starting from the state $|\downarrow\rangle|\xi\rangle$, application of the SDF ideally results in the SWES

$$|\psi(\alpha)\rangle = \frac{1}{\sqrt{2}} (|+\rangle|\alpha, \xi\rangle - |-\rangle|-\alpha, \xi\rangle) \quad (2)$$

where we use the notation $|\alpha, \xi\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle$ with the displacement operator $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$. A projective measurement of the spin performed in the $\hat{\sigma}_z$ basis gives the probability of being $|\downarrow\rangle$ as $P(\downarrow) = (1 + X)/2$, where $X = \langle\alpha, \xi| -\alpha, \xi\rangle = \langle-\alpha, \xi|\alpha, \xi\rangle$ gives the overlap between the two displaced motional states, which can be written as

$$X(\alpha, \xi) = e^{-2|\alpha|^2(\exp(2r)\cos^2(\Delta\phi) + \exp(-2r)\sin^2(\Delta\phi))} \quad (3)$$

where $\Delta\phi = \arg(\alpha) - \phi_s/2$. When $\Delta\phi = 0$, the SDF is aligned with the squeezed quadrature of the state, whereas for $\Delta\phi = \pi/2$, the SDF is

¹Institute for Quantum Electronics, Eidgenössische Technische Hochschule Zürich, Otto-Stern-Weg 1, 8093 Zürich, Switzerland. [†]Present address: Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK.

aligned with the anti-squeezed quadrature. At displacements for which X gives a measurable signal, monitoring the spin population as a function of the force duration τ for different choices of $\Delta\phi$ allows us to characterize the spatial variation of the initial squeezed wavepacket^{15–17}. For values of $|\alpha|^2$ greater than the wavepacket variance along the direction of the force, the state in equation (2) is a distinct superposition of squeezed wavepackets that have overlap close to zero and are entangled with the internal state. For $r = 0$ (no squeezing) the state reduces to the familiar SC states that have been produced in previous work^{1,3–5}. For $r > 0$, the superposed oscillator states are the displaced-squeezed states^{19,20}.

The experiments use a single trapped $^{40}\text{Ca}^+$ ion, which mechanically oscillates on its axial vibrational mode with a frequency close to $\omega_z/(2\pi) = 2.1$ MHz. This mode is well resolved from all other modes. We encode a pseudo-spin system in the internal electronic states $|\downarrow\rangle \equiv |S_{1/2}, M_J = 1/2\rangle$ and $|\uparrow\rangle \equiv |D_{5/2}, M_J = 3/2\rangle$. All coherent manipulations, including the squeezed-state preparation and the SDF, make use of the quadrupole transition between these levels at 729 nm, with a Lamb–Dicke parameter of $\eta \approx 0.05$ for the axial mode. This is small enough for the experiments to be well described using the LDA (a discussion of this approximation is given in the Methods)²¹.

We apply the SDF directly after the squeezed vacuum state has been prepared by reservoir engineering and the internal state has been prepared in $|\downarrow\rangle$ by optical pumping (in the ideal case, the ion is already in the correct state and this step has no effect). Figure 1 shows the results of measuring $\langle\hat{\sigma}_z\rangle$ after applying displacements along the two principal axes of the squeezed state alongside the same measurement made using an ion prepared in the motional ground state. To extract relevant parameters regarding the SDF and the squeezing, we fit the data using $P(\downarrow) = (A + B X(\alpha, \xi))/2$, where the parameters A and B account for experimental imperfections such as shot-to-shot fluctuations

in the magnetic field (Methods). Fitting the ground-state data with r fixed to zero allows us to extract $\Omega/(2\pi) = 13.25 \pm 0.40$ kHz (here and in the rest of the paper, all errors are given as s.e.m.). We then fix this when performing independent fits to the squeezed-state data for $\Delta\phi = 0$ and $\Delta\phi = \pi/2$. Each of these fits allows us to extract an estimate for the squeezing parameter r . For both the squeezed and anti-squeezed quadratures we obtain consistent values with a mean of $r = 1.08 \pm 0.03$, which for a pure state would correspond to a 9.4 dB reduction in the variance of the squeezed quadrature. The inset of Fig. 1 shows the spin population as a function of the SDF phase ϕ_D with the SDF duration fixed to 20 μs . This is also fitted using the same equation described above, and we obtain $r = 1.13 \pm 0.03$.

The loss of overlap between the two wavepackets indicates that a SWES has been created. To verify that these states are coherent superpositions, we recombine the wavepackets by applying a second ‘return’ SDF pulse for which the phase of both the red and blue sideband laser frequency components is shifted by π relative to the first. This reverses the direction of the force applied to the motional states for both the $|+\rangle$ and $|-\rangle$ spin states. In the ideal case a state displaced to $\alpha(\tau_1)$ by a first SDF pulse of duration τ_1 has a final displacement of $\delta\alpha = \alpha(\tau_1) - \alpha(\tau_2)$ after the return pulse of duration τ_2 . For $\tau_1 = \tau_2$, $\delta\alpha = 0$ and the measured probability of finding the spin state in $|\downarrow\rangle$ is 1. In the presence of decoherence and imperfect control, the probability with which the ion returns to the $|\downarrow\rangle$ state will be reduced. In Fig. 2 we show revivals in the spin coherence for the same initial squeezed vacuum state as was used for the data in Fig. 1. The data include a range of different τ_1 . For the data for which the force was applied along the squeezed axis of the state ($\Delta\phi = 0$), partial revival of the coherence is observed for SDF durations up to 250 μs . For $\tau_1 = 250$ μs the maximum separation of the two distinct oscillator wavepackets is $|\Delta\alpha| > 19$, which is 56 times the r.m.s. width of the squeezed wavepacket in phase space. The amplitude of revival of this state is similar to what we observe when applying the SDF to a ground-state cooled ion. The loss of coherence as a function of the displacement duration is consistent with the effects of magnetic-field-induced spin dephasing and motional heating^{14,22}. When the force is applied along the anti-squeezed quadrature ($\Delta\phi = \pi/2$), we observe that the strength of the revival decays more rapidly than for displacements with $\Delta\phi = 0$. Simulations of the dynamics using a quantum Monte Carlo wavefunction approach including sampling over a magnetic field distribution indicate that this is caused by shot-to-shot fluctuations of the magnetic field (Methods).

We are also able to monitor the number-state distributions of the motional wavepackets as a function of the duration of the SDF. This provides a second measurement of the parameters of the SDF and the initial squeezed wavepacket, which has similarities with the homodyne measurement used in optics^{23,24}. To do this, we optically pump the spin state into $|\downarrow\rangle$ after applying the SDF. This procedure destroys the phase relationship between the two motional wavepackets, resulting in the mixed oscillator state $\hat{\rho}_{\text{mixed}} = (|\alpha, \xi\rangle\langle\alpha, \xi| + |-\alpha, \xi\rangle\langle-\alpha, \xi|)/2$ (we estimate that the photon recoil during optical pumping results in a decrease in the fidelity of our experimental state relative to $\hat{\rho}_{\text{mixed}}$ by $< 3\%$, which would not be observable in our measurements). The two parts of this mixture have the same number-state distribution, which is that of a displaced-squeezed state^{19,20}. To extract this distribution, we drive Rabi oscillations on the blue-sideband transition²⁵ and monitor the subsequent spin population in the $\hat{\sigma}_z$ basis. Figure 3 shows this evolution for SDF durations of $\tau = 0, 30, 60$ and 120 μs . For $\tau = 30$ and 60 μs , the results from displacements applied parallel to the two principal axes of the squeezed state are shown ($\Delta\phi = 0$ and $\pi/2$). We obtain the number-state probability distribution $p(n)$ from the spin state population by fitting the data using a form $P(\downarrow) = bt + \frac{1}{2} \sum_n p(n)(1 + e^{-\gamma t} \cos(\Omega_{n,n+1}t))$, where t is the blue-sideband pulse duration, $\Omega_{n,n+1}$ is the Rabi frequency for the transition between the $|\downarrow\rangle|n\rangle$ and $|\uparrow\rangle|n+1\rangle$ states, and γ is a phenomenological decay

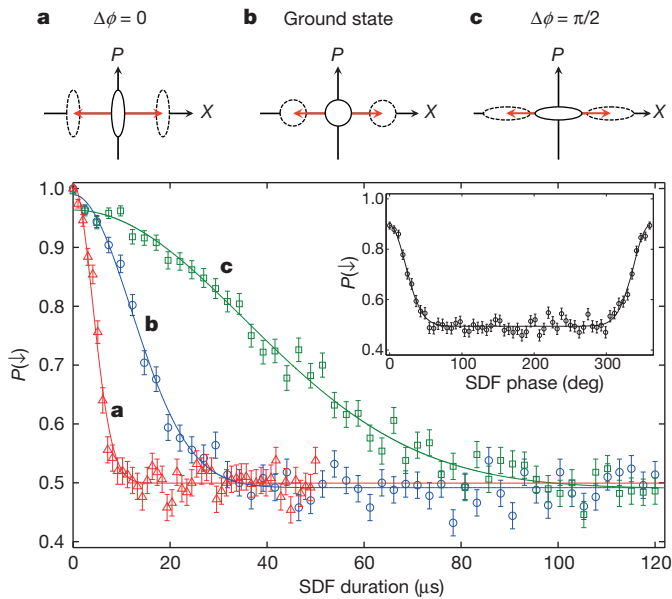


Figure 1 | Spin population evolution due to spin–motion entanglement. Projective measurement of the spin in the $\hat{\sigma}_z$ basis as a function of SDF duration. **a**, Forces parallel to the squeezed quadrature (red triangles). **b**, An ion initially prepared in the motional ground state (blue circles). **c**, Forces parallel to the anti-squeezed quadrature (green squares). The inset shows a scan of the phase of the SDF for an initial squeezed state with the force duration fixed at 20 μs . Each data point is the result of > 300 repetitions of the experimental sequence. Results are shown as means \pm s.e.m.; the error bars were generated under the assumption that the dominant source of fluctuations was quantum projection noise.

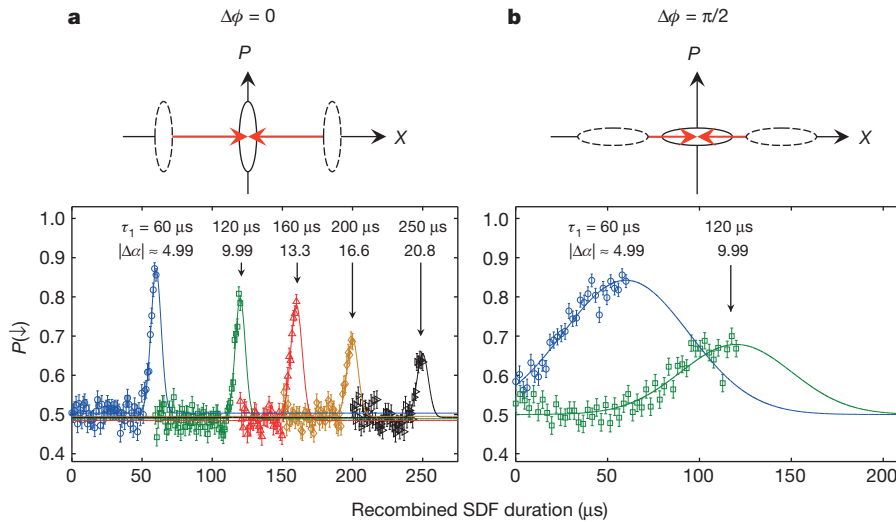


Figure 2 | Revival of the spin coherence. Spin populations as a function of the duration of the second SDF pulse with the spin phase shifted by π relative to the first pulse. **a**, Forces parallel to the squeezed quadrature. **b**, Forces parallel to the anti-squeezed quadrature. In all cases an increase in the spin population is seen at the time when the two motional states are overlapped, which corresponds to the time τ_1 used for the first SDF pulse. The value of τ_1 and the corresponding $|\Delta\alpha|$ calculated from the measured Rabi frequency are written above the revival of each data set. The fractional error on the mean of each of the estimated $|\Delta\alpha|$ is about 3%. The solid lines are fitted curves using the same form as for the fits in Fig. 1 with the overlap function $X(\delta x, \xi)$. The values of r obtained are consistent with the data in Fig. 1. Results are shown as means \pm s.e.m.; the error bars were generated under the assumption that the dominant source of fluctuations was quantum projection noise.

parameter^{25,26}. The parameter b accounts for gradual pumping of population into the state $|\uparrow\rangle|0\rangle$ due to frequency noise on our laser^{18,27}. It is negligible when $p(0)$ is small. The resulting $p(n)$ are then fitted using the theoretical form for the displaced-squeezed states (Methods). The number-state distributions show a clear dependence on the phase of the force, which is also reflected in the spin population evolution. Figure 4 shows

the Mandel Q parameters of the experimentally obtained number-state distributions, defined as $Q = \langle (\Delta n)^2 \rangle / \langle n \rangle - 1$, in which $\langle (\Delta n)^2 \rangle$ and $\langle n \rangle$ are the variance and mean of $p(n)$, respectively²⁸. The solid lines are the theoretical curves given in ref. 19 for $r = 1.08$, and are in agreement with our experimental results. For displacements along the short axis of the squeezed state (Fig. 3), the collapse and revival behaviour of the time

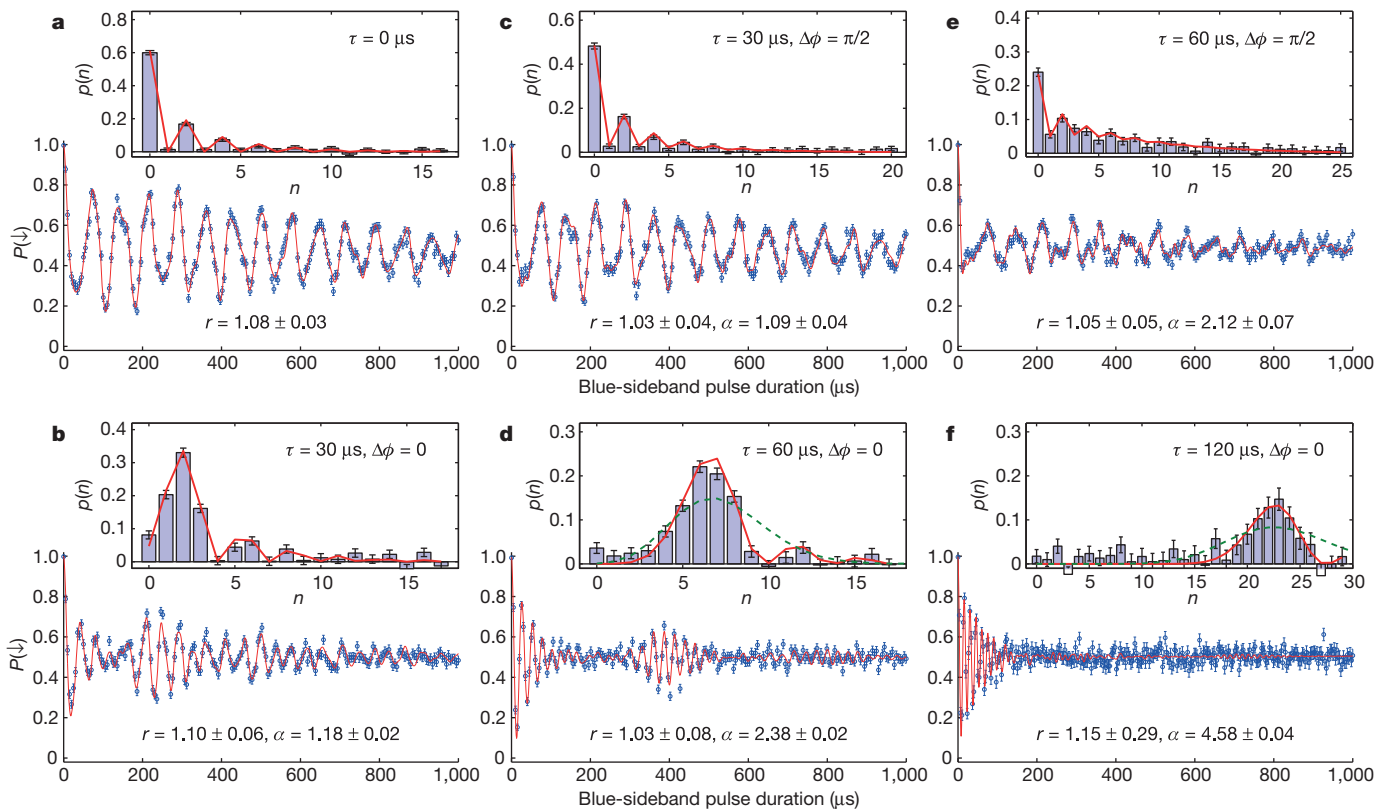


Figure 3 | Evolution of displaced-squeezed-state mixtures. The observed blue-sideband oscillations and the corresponding number-state probability distributions for the SDF applied along the two principal axes of the squeezed state and with different durations. **a**, Initial squeezed vacuum state. **b, d, f**, Forces parallel to the squeezed quadrature. **c, e**, Forces parallel to the anti-squeezed quadrature. For $\tau = 30 \mu\text{s}$ the obtained parameters are consistent within statistical errors. For $\tau = 60 \mu\text{s}$ the displacement along the anti-squeezed quadrature (**e**) results in a large spread in the number-state probability

distribution, with the result that in the fitting r and α are positively correlated; the errors stated do not take account of this. We think that this accounts for the apparent discrepancy between the values of r and α obtained for $\tau = 60 \mu\text{s}$. The dashed green line in the insets of **d** and **f** is the Poisson distribution for the same $\langle n \rangle$ as the created displaced-squeezed-state mixture, which is given by $\langle n \rangle = |\alpha|^2 + \sinh^2 r$ (ref. 19). Results are shown as means \pm s.e.m.; the error bars were generated under the assumption that the dominant source of fluctuations was quantum projection noise.

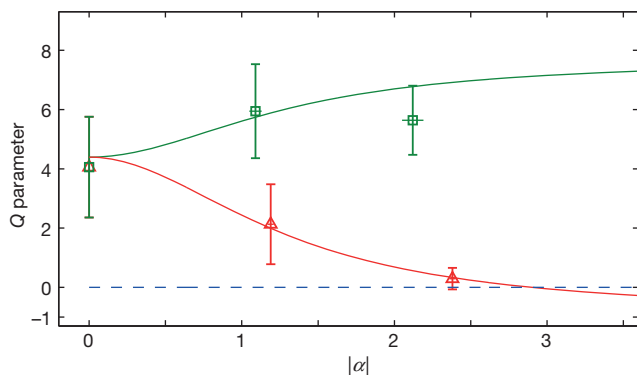


Figure 4 | Mandel Q parameter for the displaced-squeezed states. Results for displacements along the squeezed quadrature (red triangles) and the anti-squeezed quadrature (green squares). All values are calculated from the experimental data given in Fig. 3, taking the propagation of error into account. The solid lines are theoretical curves for displacements along the squeezed (red) and anti-squeezed (green) quadratures of an initial state with $r = 1.08$. The values of $|\alpha|$ are obtained from fits to the respective $p(n)$ (Fig. 3), with error bars comparable to the size of the symbol. The point at $|\alpha| = 0$ is the squeezed vacuum state.

evolution of $P(\downarrow)$ is reminiscent of the Jaynes–Cummings Hamiltonian applied to a coherent state²⁹, but it has more oscillations before the ‘collapse’ for a state of the same $\langle n \rangle$. This is surprising because the statistics of the state is not sub-Poissonian. We attribute this to the fact that this distribution is more peaked than that of a coherent state with the same $\langle n \rangle$, which is obvious when the two distributions are plotted over one another (Fig. 3d, f). The increased variance of the squeezed state then arises from the extra populations at high n , which are too small to make a visible contribution to the Rabi oscillations. For the squeezing parameter in our experiments, sub-Poissonian statistics would be observed only for $|\alpha| > 3$. For $\tau = 120 \mu\text{s}$ we obtain a consistent value of r and $|\alpha| = 4.6$ only in the case in which we include a fit parameter for scaling of the theoretical probability distribution, obtaining a fitted scaling of 0.81 ± 0.10 (Methods). The reconstruction of the number-state distribution is incomplete because we cannot extract populations with $n > 29$ as a result of frequency crowding in the $\sqrt{n+1}$ dependence of the Jaynes–Cummings dynamics. We therefore do not include these results in Fig. 4. Measurement techniques made in a squeezed-state basis¹⁸ could avoid this problem; however, these are beyond our current experimental capabilities for states of this size.

We have generated entangled superposition states between the internal and motional states of a single trapped ion in which the superposed motional wavepackets are of a squeezed Gaussian form. These states present new possibilities both for metrology and for continuous variable quantum information. In an interferometer based on SC states separated by $|\Delta x|$, the interference contrast depends on the final overlap of the recombined wavepackets. Fluctuations in the frequency of the oscillator result in a reduced overlap, but this effect can be improved by a factor $\exp[-|\Delta x|^2(e^{-2r}-1)/2]$ if the wavepackets are squeezed in the same direction as the state separation (Methods). In quantum information with continuous variables, the computational basis states are distinguishable because they are separated in phase space by $|\Delta x|$ and thus do not overlap^{7–9}. The decoherence times of such superpositions typically scale as $1/|\Delta x|^2$ (ref. 22). The use of states squeezed along the displacement direction reduces the required displacement for a given overlap by e^r , increasing the resulting coherence time by e^{2r} , which is a factor of 9 in our experiments. We therefore expect these states to open up new possibilities for quantum-state engineering and control.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 December 2014; accepted 26 March 2015.

- Wineland, D. J. Nobel lecture. Superposition, entanglement, and raising Schrödinger’s cat. *Rev. Mod. Phys.* **85**, 1103–1114 (2013).
- Haroche, S. Nobel lecture. Controlling photons in a box and exploring the quantum to classical boundary. *Rev. Mod. Phys.* **85**, 1083–1102 (2013).
- Monroe, C., Meekhof, D. M., King, B. E. & Wineland, D. J. A ‘Schrödinger cat’ superposition state of an atom. *Science* **272**, 1131–1136 (1996).
- McDonnell, M. J. *et al.* Long-lived mesoscopic entanglement outside the Lamb–Dicke regime. *Phys. Rev. Lett.* **98**, 063603 (2007).
- Haljan, P. C., Brickman, K.-A., Deslauriers, L., Lee, P. J. & Monroe, C. Spin-dependent forces on trapped ions for phase-stable quantum gates and motional Schrödinger cat states. *Phys. Rev. Lett.* **94**, 153602 (2005).
- Munro, W. J., Nemoto, K., Milburn, G. J. & Braunstein, S. L. Weak-force detection with superposed coherent states. *Phys. Rev. A* **66**, 023819 (2002).
- Weedbrook, C. *et al.* Gaussian quantum information. *Rev. Mod. Phys.* **84**, 621–669 (2012).
- Gottesman, D., Kitaev, A. & Preskill, J. Encoding a qubit in an oscillator. *Phys. Rev. A* **64**, 012310 (2001).
- Bartlett, S. D., de Guise, H. & Sanders, B. C. Quantum encodings in spin systems and harmonic oscillators. *Phys. Rev. A* **65**, 052316 (2002).
- Monroe, C., Meekhof, D. M., King, B. E., Itano, W. M. & Wineland, D. J. Demonstration of a fundamental quantum logic gate. *Phys. Rev. Lett.* **75**, 4714–4717 (1995).
- Collaboration, L. S. *et al.* A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nature Phys.* **7**, 962–965 (2011).
- Aasi, J. *et al.* Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light. *Nature Photon.* **7**, 613–619 (2013).
- Vlastakis, B. *et al.* Deterministically encoding quantum information using 100-photon Schrödinger’s cat states. *Science* **342**, 607–610 (2013).
- Hempel, C. *et al.* Entanglement-enhanced detection of single-photon scattering events. *Nature Photon.* **7**, 630–633 (2013).
- Hao-Sheng, Z., Ai-Qin, H., Qiong, L. & Le-Man, K. Direct measurement of squeezing in the motion of trapped ions. *Chin. Phys. Lett.* **22**, 798–800 (2005).
- Gerritsma, R. *et al.* Quantum simulation of the Dirac equation. *Nature* **463**, 68–71 (2010).
- Casanova, J., López, C. E., García-Ripoll, J. J., Roos, C. F. & Solano, E. Quantum tomography in position and momentum space. *Eur. Phys. J. D* **66**, 1–5 (2012).
- Kienzler, D. *et al.* Quantum harmonic oscillator state synthesis by reservoir engineering. *Science* **347**, 53–56 (2015).
- Caves, C. M. Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **23**, 1693–1708 (1981).
- Yuen, H. P. Two-photon coherent states of the radiation field. *Phys. Rev. A* **13**, 2226–2243 (1976).
- Wineland, D. J. *et al.* Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl. Inst. Stand. Technol.* **103**, 259–328 (1998).
- Turchette, Q. A. *et al.* Decoherence and decay of motional quantum states of a trapped atom coupled to engineered reservoirs. *Phys. Rev. A* **62**, 053807 (2000).
- Breitenbach, G., Schiller, S. & Mlynek, J. Measurement of the quantum states of squeezed light. *Nature* **387**, 471–475 (1997).
- Ourjoumtsev, A., Jeong, H., Tualle-Brouiri, R. & Grangier, P. Generation of optical ‘Schrödinger cats’ from photon number states. *Nature* **448**, 784–786 (2007).
- Meekhof, D. M., Monroe, C., King, B. E., Itano, W. M. & Wineland, D. J. Generation of nonclassical motional states of a trapped atom. *Phys. Rev. Lett.* **76**, 1796–1799 (1996). Erratum. *Phys. Rev. Lett.* **77**, 2346 (1996).
- Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
- Di Fidio, C. & Vogel, W. Damped Rabi oscillations of a cold trapped ion. *Phys. Rev. A* **62**, 031802 (2000).
- Mandel, L. Sub-Poissonian photon statistics in resonance fluorescence. *Opt. Lett.* **4**, 205–207 (1979).
- Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms and Cavities and Photons* (Oxford Univ. Press (2006))

Acknowledgements We thank J. Alonso and F. Leupold for comments on the manuscript, and F. Leupold, F. Lindenfelser, J. Alonso, M. Sepiol, K. Fisher and C. Flühmann for contributions to the experimental apparatus. We acknowledge support from the Swiss National Science Foundation under grant number 200021 134776, and through the National Centre of Competence in Research for Quantum Science and Technology (QSIT).

Author Contributions Experimental data were taken by H.-Y.L., D.K. and L.d.C., using an apparatus primarily built by D.K., H.-Y.L. and B.C.K., and with significant contributions from L.d.C., V.N. and M.M. Data analysis was performed by H.-Y.L. and J.P.H. The paper was written by J.P.H. and H.-Y.L., with input from all authors. The study was conceived by J.P.H.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.-Y.L. (hylo@phys.ethz.ch) or J.P.H. (jhome@phys.ethz.ch).

METHODS

Experimental details. The experiments make use of a segmented linear Paul trap with an ion–electrode distance of $\sim 185 \mu\text{m}$. Motional heating rates from the ground state for a calcium ion in this trap have been measured to be 10 ± 1 quanta s^{-1} , and the coherence time for the number-state superposition $(|0\rangle + |1\rangle)/\sqrt{2}$ has been measured to be 32 ± 3 ms.

The first step of each experimental run involves cooling all modes of motion of the ion close to the Doppler limit by using laser light at 397 and 866 nm. The laser beam used for coherent control of the two-level pseudo-spin system addresses the narrow-linewidth transition $|\downarrow\rangle \equiv |S_{1/2}, M_J = 1/2\rangle \leftrightarrow |\uparrow\rangle \equiv |D_{5/2}, M_J = 3/2\rangle$ at 729 nm. This transition is resolved by 200 MHz from all other internal state transitions in the applied magnetic field of 119.6 G. The SDFs and the reservoir engineering¹⁸ in our experiment require the application of a bichromatic light field. We generate both frequency components with the use of acousto-optic modulators (AOMs) starting from a single laser stabilized to an ultra-high-finesse optical cavity with a resulting linewidth of < 600 Hz (at which point magnetic field fluctuations limit the qubit coherence). We apply pulses of 729 nm laser light with a double-pass AOM to which we apply a single radiofrequency tone, followed by a single-pass AOM to which two radiofrequency tones are applied. After this second AOM, both frequency components are coupled into the same single-mode fibre before delivery to the ion. The double-pass AOM is used to switch the light on and off. Optical pumping to $|\downarrow\rangle$ is implemented using a combination of linearly polarized light fields at 854, 397 and 866 nm. The internal state of the ion is read out by state-dependent fluorescence using laser fields at 397 and 866 nm.

The 729 nm laser beam enters the trap at 45° to the z axis of the trap, resulting in a Lamb–Dicke parameter of $\eta \approx 0.05$ for the axial mode. For this Lamb–Dicke parameter, we have verified whether for displacements up to $|\alpha| = 9.75$ the dynamics can be well described with the LDA. We simulate the wavepacket dynamics by using the interaction Hamiltonian with and without LDA. In the simulation we apply the SDF to an ion prepared in $|\downarrow\rangle|\xi\rangle$. The interaction Hamiltonian for a single trapped ion coupled to a single-frequency laser field can be written as²⁶

$$\hat{H}_I = \frac{\hbar}{2} \Omega_0 \hat{\sigma}_+ \exp\{i\eta(\hat{a}e^{-i\omega_z t} + \hat{a}^\dagger e^{i\omega_z t})\} e^{i(\phi - \delta t)} + \text{H.c.}$$

where Ω_0 is the interaction strength, $\hat{\sigma}_+ = |\uparrow\rangle\langle\downarrow|$, \hat{a} and \hat{a}^\dagger are motional annihilation and creation operators, ω_z is the vibrational frequency of the ion, ϕ is the phase of the laser, $\delta = \omega_l - \omega_a$ is the detuning of the laser from the atomic transition, and ‘H.c.’ is the Hermitian conjugate of the first term. In the laboratory, the application of the SDF involves simultaneously driving both the blue-sideband and red-sideband transitions resonantly, resulting in the Hamiltonian $\hat{H}_{\text{tot}} = \hat{H}_{\text{bsb}} + \hat{H}_{\text{rsb}}$, where $\delta = \omega_z$ in \hat{H}_{bsb} and $\delta = -\omega_z$ in \hat{H}_{rsb} . Starting from $|\downarrow\rangle|\xi\rangle$, the evolution of the state cannot be solved analytically. We perform a numerical simulation in which we retain only the resonant terms in the Hamiltonian. Extended Data Fig. 1 shows the quasi-probability distributions in phase space for chosen values of the SDF duration τ . These are compared with results obtained using the LDA. For $\tau = 60 \mu\text{s}$ both cases are similar, resulting in $|\alpha| \approx 2.4$. For $\tau = 250 \mu\text{s}$ the squeezed-state wavepackets are slightly distorted and the displacement is 4% smaller for the full simulation than for the LDA form. Considering the levels of error arising from imperfect control and decoherence for forces of this duration, we do not consider this effect to be significant in our experiments.

Simulations for the coherence of SWESs. After creating SWESs, we deduce that coherence is retained throughout the creation of the state by applying a second SDF pulse to the ion, which recombines the two separated wavepackets and disentangles the spin from the motion. The revival in the spin coherence is not perfect, because of decoherence and imperfect control in the experiment. One dominant source causing decoherence of the superpositions is spin decoherence due to magnetic field fluctuations. We have performed quantum Monte Carlo wavefunction simulations to investigate the coherence of the SWES in the presence of such a decoherence mechanism. We simulate the effect of a sinusoidal fluctuation of the magnetic field on a timescale that is long compared with the duration of the coherent control sequence, which is consistent with the noise that we observe on our magnetic field coil supply (at 10 and 110 Hz) and from ambient fluctuations due to electronics equipment in the room. The amplitude of these fluctuations is set to 2.2 mG, giving rise to the spin coherence time of 180 μs , which we measured using Ramsey experiments on the spin alone. Because the frequency of fluctuations is slow compared with the sequence length, we fix the field for each run of the simulation but sample its value from a probability distribution derived from a sinusoidal oscillation. In Extended Data Fig. 2 we show the effect of a single shot taken at a fixed qubit-oscillator detuning of 1.5 kHz, and in Extended Data Fig. 3 we show the average over the distribution. In both figures, results are shown for the SDF applied along the two principal axes of the squeezed vacuum state as well as for the motional ground state using force durations of 60 and 120 μs . We also show the results of applying the second SDF pulse, resulting in partial revival of the spin coherence. It can be seen that when the SDF is applied along the

anti-squeezed quadrature, the strength of the revival decays more rapidly, and $P(1)$ oscillates around 0.5. This effect can be seen in the data shown in Fig. 2.

Number-state probability distributions for the displaced-squeezed state. For Fig. 3 we characterize the probability distribution for the number states of the oscillator. This is performed by driving the blue-sideband transition $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n+1\rangle$ and fitting the obtained spin population evolution using

$$P(\downarrow) = bt + \frac{1}{2} \sum_n p(n)(1 + e^{-\gamma t} \cos(\Omega_{n,n+1}t)) \quad (4)$$

where t is the blue-sideband pulse duration, $p(n)$ are the number-state probabilities for the motional state we are concerned with, and γ is an empirical decay parameter^{25,26}. In the results presented here we do not scale this decay parameter with n as was done in ref. 25. We also fitted the data including such a scaling and saw consistent results. The Rabi frequency coupling $|\downarrow\rangle|n\rangle$ to $|\uparrow\rangle|n+1\rangle$ is $\Omega_{n,n+1} = \Omega_0 \langle n | e^{i\eta(\hat{a}^\dagger + \hat{a})} | n+1 \rangle = \Omega_0 e^{-\eta^2/2} \eta L_n^1(\eta^2)/\sqrt{n+1}$. For small n , this scales as $\sqrt{n+1}$, but because the states include significant populations at higher n we use the complete form including the generalized Laguerre polynomial $L_n^1(x)$. The parameter b in the first term accounts for a gradual pumping of population into the state $|\uparrow\rangle|0\rangle$, which is not involved in the dynamics of the blue-sideband pulse^{18,27}. This effect is negligible when $p(0)$ is small.

After extracting $p(n)$ from $P(\downarrow)$, we fit it using the number-state probability distribution for the displaced-squeezed state³⁰:

$$p(n) = \kappa \frac{(\frac{1}{2} \tanh r)^n}{n! \cosh r} \exp \left[-|\alpha|^2 - \frac{1}{2} (\alpha^{*2} e^{i\phi_s} + \alpha^2 e^{-i\phi_s}) \tanh r \right] \left| H_n \left(\frac{\alpha \cosh r + \alpha^* e^{i\phi_s} \sinh r}{\sqrt{e^{i\phi_s} \sinh 2r}} \right) \right|^2$$

where κ is a constant that accounts for the infidelity of the state during the application of SDF, and the $H_n(x)$ are the Hermite polynomials. The direction of the SDF is aligned along either the squeezing quadrature or the anti-squeezing quadrature of the state. Therefore we set $\arg(\alpha) = 0$ and fix $\phi_s = 0$ and π for fitting the data of the short axis and the long axis of the squeezed state, respectively. This allows us to obtain the values of r and $|\alpha|$ for the state we created. For the cases of smaller displacements (from Fig. 3a–e), we set $\kappa = 1$. For the data set of $|\alpha| \approx 4.6$ (Fig. 3f), κ is a fitting parameter that gives us a value of 0.81 ± 0.1 . We note that in this case 4% of the expected population lies above $n = 29$ but we are unable to extract these populations from our data.

The Mandel Q parameter²⁸ is defined as

$$Q = \frac{\langle (\Delta n)^2 \rangle - \langle n \rangle}{\langle n \rangle}$$

where $\langle n \rangle$ and $\langle (\Delta n)^2 \rangle$ are the mean and variance of the probability distribution. For a displaced-squeezed state these are given in ref. 19 as

$$\langle (\Delta n)^2 \rangle = |\alpha \cosh r - \alpha^* e^{i\phi_s} \sinh r|^2 + 2 \cosh^2 r \sinh^2 r$$

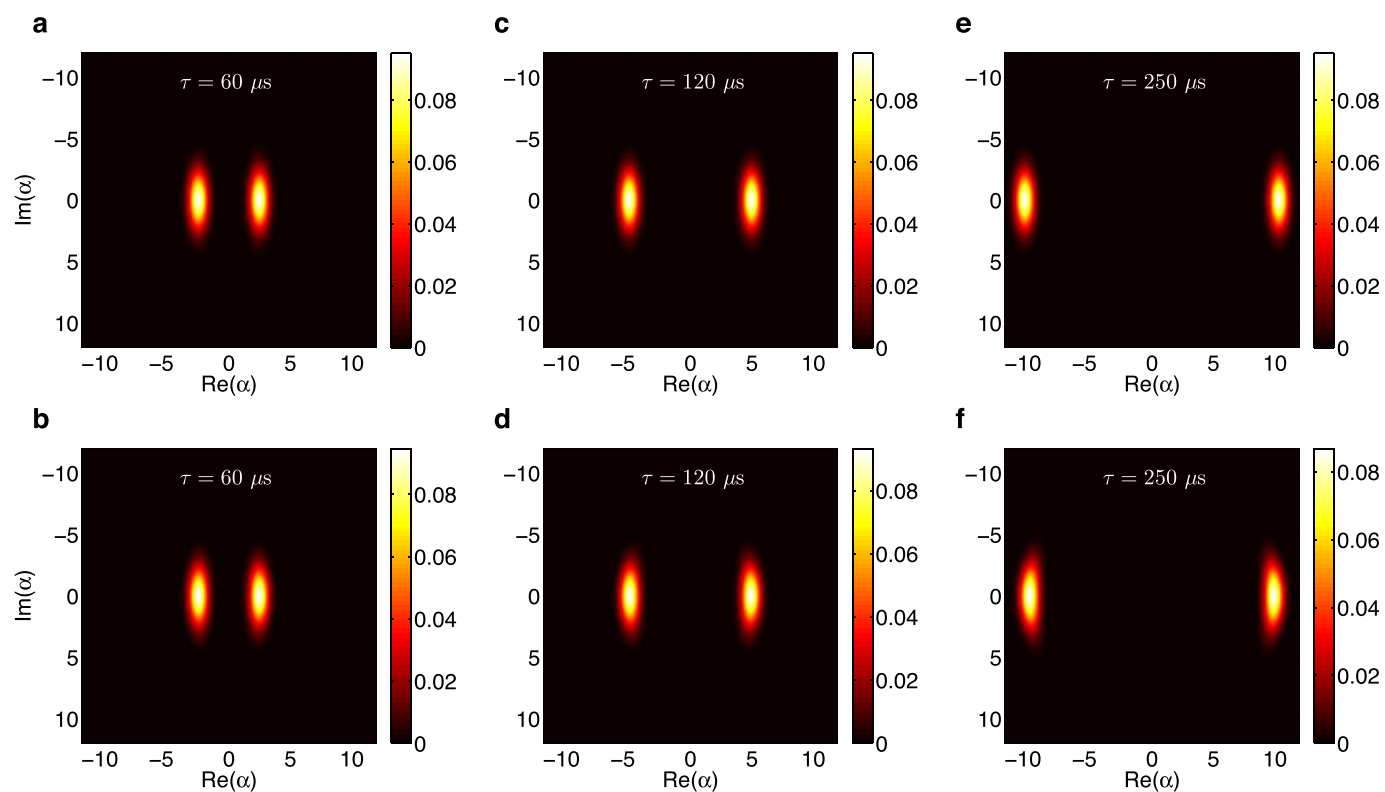
$$\langle n \rangle = |\alpha|^2 + \sinh^2 r$$

These forms were used to produce the curves given in Fig. 4.

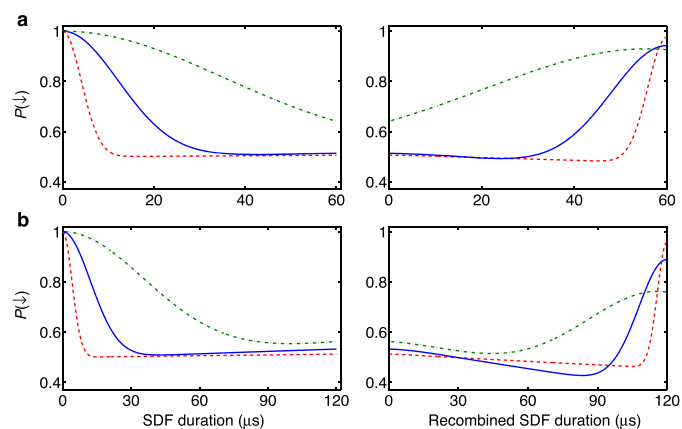
Applications of SWESs. The SWES may offer new possibilities for sensitive measurements that are robust against certain types of noise. An example is illustrated in Extended Data Fig. 4, in which we compare an interferometry experiment involving the use of a SWES versus a more standard SC state based on coherent states. In both cases the superposed states have a separation of $|2\alpha|$ obtained using a SDF. For the SWES this force is aligned along the squeezed quadrature of the state. The interferometer is closed by inverting the initial SDF, resulting in a residual displacement that in the ideal case is zero. One form of noise involves shot-to-shot fluctuations in the oscillator frequency. On each run of the experiment, this would result in a small phase shift $\Delta\theta$ arising between the two superposed motional states. As a result, after the application of the second SDF pulse the residual displacement would be $\alpha_R = 2i\alpha \sin(\Delta\theta/2)$, which corresponds to the states being separated along the P axis in the rotating-frame phase space. The final state of the system would then be $|\psi(\alpha_R)\rangle$ with a corresponding state overlap given by $X(\alpha_R, \xi)$. Therefore the contrast will be higher for the SWES (Extended Data Fig. 4a) than for the coherent SC state (Extended Data Fig. 4b) by a factor

$$\exp[-2|\alpha_R|^2(e^{-2r} - 1)]$$

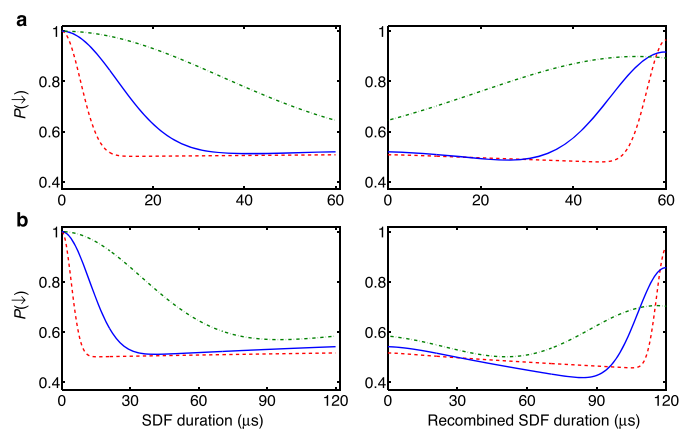
Although in our experiments other sources of noise dominate, in other systems such oscillator dephasing may be more significant.



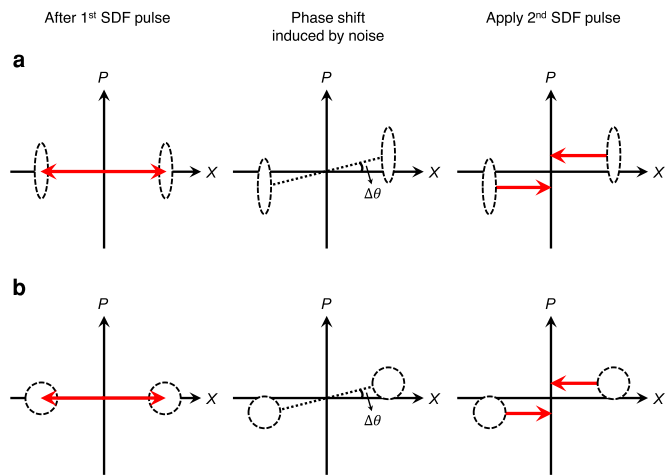
Extended Data Figure 1 | Quasi-probability distributions for displaced-squeezed states in phase space using LDA and non-LDA. **a, c, e,** The simulation results using LDA with different SDF durations. **b, d, f,** The results simulated using the full Hamiltonian.



Extended Data Figure 2 | Coherence of cat states with fixed magnetic field noise. The magnetic-field-induced energy-level shift of 1.5 kHz is used in this simulation. **a**, The duration of both SDF pulses is 60 μs . **b**, The duration of both SDF pulses is 120 μs . Dashed red and dash-dot green curves show the SDF aligned along the squeezed and anti-squeezed quadratures. The blue trace is for the SDF applied to a ground-state cooled ion.



Extended Data Figure 3 | Coherence of cat states with a magnetic field fluctuation distribution. With the assumption that the magnetic field exhibits a 50 Hz sinusoidal pattern with an amplitude of 2.2 mG, this plot shows the simulation results by taking an average over 100 samples on the field distribution. **a**, The duration of both SDF pulses is 60 μ s. **b**, The duration of both SDF pulses is 120 μ s. Definitions of the curve specification are the same as in Extended Data Fig. 2.



Extended Data Figure 4 | Possible application of using SWESs for interferometry. **a**, Use of squeezed-state wavepackets. **b**, Use of ground-state wavepackets. The first SDF pulse is used to create a spin–motion-entangled state. In the middle, a small phase shift $\Delta\theta$ is induced by shot-to-shot fluctuation in the oscillator frequency before the application of the second SDF pulse, which recombines the two distinct oscillator wavepackets.

Non-Joulian magnetostriction

Harsh Deep Chopra¹ & Manfred Wuttig²

All magnets elongate and contract anisotropically when placed in a magnetic field, an effect referred to as Joule magnetostriction¹. The hallmark of Joulian magnetostriction is volume conservation², which is a broader definition applicable to self-accommodation of ferromagnetic, ferroelectric or ferroelastic domains in all functional materials^{3–10}. Here we report the discovery of ‘giant’ non-volume-conserving or non-Joulian magnetostriction (NJM). Whereas Joulian strain is caused by magnetization rotation, NJM is caused by facile (low-field) reorientation of magnetoelastically and magnetostatically autarkic (self-sufficient) rigid micro-‘cells’, which define the adaptive structure, the origin of which is proposed to be elastic gradients ultimately caused by charge/spin density waves^{11–13}. The equilibrium adaptive cellular structure is responsible for long-sought non-dissipative (hysteresis-free), linearly reversible and isotropic magnetization curves along all directions within a single crystal. Recently discovered Fe-based high magnetostriction alloys^{14,15} with special thermal history are identified as the first members of this newly discovered magnetic class. The NJM paradigm provides consistent interpretations of seemingly confounding properties of Fe-based alloys, offers recipes to develop new highly magnetostrictive materials, and permits simultaneously large actuation in longitudinal and transverse directions without the need for stacked composites.

Figures 1, 2 and 3 represent the core of our experimental results; Fig. 1 shows NJM of two Fe–Ga single crystals with widely different compositions and thermal treatments; Fig. 2 shows the unusual characteristics of isotropic and non-hysteretic magnetization curves in Fe–Ga, Fe–Al and Fe–Ge single crystals; and Fig. 3 shows the cellular microstructure of Fe–Ga central to the interpretation of our results. Figure 4 summarizes the rule of mixture governing the magnitude of NJM in various directions.

Single crystals of Fe–Ga, Fe–Ge and Fe–Al were in the form of circular disks ~5 mm in diameter and 0.4–0.5-mm thick, with [001] direction normal to the disk plane. Magnetization curves were measured at room temperature using a vibrating sample magnetometer. The magnetostriction measurements were made by the standard strain gauge method. The magnetic domain structure was studied using the high-resolution interference contrast colloid (ICC) method, which has been described in detail previously^{7,16}. See Methods for additional details.

The characteristic features of NJM are illustrated in Fig. 1a and b for Fe_{73.9}–Ga_{26.1} and Fe_{82.9}–Ga_{17.1} single crystals, respectively. Both alloys produce a large longitudinal magnetostriction expansion strain of ~200 p.p.m. when the applied field is along the [100] direction, and ~100 p.p.m. of longitudinal strain when the field is along the [110] direction. Measurement of magnetostriction along various directions shows in fields $H \parallel [100]$, the angular dependence of the resulting deformation is positive in all directions, as shown in Fig. 1c for the Fe_{73.9}–Ga_{26.1} alloy. In other words, the sample expands in all directions and increases its volume, that is, displays NJM; a negligible magnetization normal to the plane of the disk at comparable fields implies that no volume change occurs along [001]. Figure 1c shows that the deformed state of the initially circular disk possesses uniaxial sym-

metry. The transverse strain along [010] is small but positive (16 p.p.m.) and the simultaneous expansion along [110]-type directions equals 100 p.p.m. If, on the other hand, $H \parallel [110]$, experiments again yield net expansion in all directions but now the deformation displays four-fold symmetry (Fig. 1d). In this case, both the longitudinal and transverse strains are ~110–120 p.p.m., and strain along [100] equals 130 p.p.m. The magnetostriction in Fig. 1c, d is shown highly exaggerated relative to the initial circular shape but the relative dimensional changes are to scale. Various longitudinal and transverse magnetostriction curves in the plane of the disk are shown in Extended Data Fig. 1, which was used to construct Fig. 1c, d. Although not shown, similar behaviour is observed for the Fe_{82.9}–Ga_{17.1} alloy in Fig. 1b. Magnetostriction curves were highly reproducible over repeated measurements (Extended Data Fig. 2, corresponding to various curves shown in Fig. 1a and Extended Data Fig. 1). In addition, a subset of curves was independently measured in our respective laboratories, yielding similar results. The magnitude of longitudinal magnetostriction constitutes the second anomaly; its maxima occur along the easy directions, [100] and [010], and not along the hard directions of the [110] type (Fig. 1a, b), just the opposite to conventional ferromagnets. This unusual behaviour is explained on the basis of rule of mixture for cells whose orientation can only be directed along [100] and [010] (Fig. 4). The observation of NJM is sensitive to previous heat treatment. For instance, when the temperature of a crystal whose composition was similar to the one shown in Fig. 1a was lowered from 1,033 K to room temperature in a furnace (instead of fast cooling), its NJM characteristics degraded (Extended Data Fig. 3).

The magnetization curves of these bcc Fe-rich non-Joulian magnets and the similar solid solutions of Fe–Ge and Fe–Al (Fig. 2a–c) exhibit equally unusual combinations of magnetic characteristics—they are reversible, linear, non-hysteretic curves resembling rotation, but at the same time they are identical in all crystallographic orientations, as if the material were somehow amorphous. (For reference, in an isotropic ferromagnet, $\lambda_{\parallel} = -2\lambda_{\perp}$, that is, $\lambda_{\parallel} + 2\lambda_{\perp} = 0$ or zero volume variation for Joulian magnetostriction (JM); whereas $\lambda_{\parallel} + 2\lambda_{\perp} \neq 0$ characterizes NJM.) They saturate at relatively low fields and feature unusually low anisotropy energies, as low as $\sim 10^2$ – 10^3 J m^{−3} (ref. 17). For comparison, conventional magnets such as Fe, Co or Ni display hysteretic (square) and linear magnetization curves along easy and hard axes, as shown schematically in Fig. 2d, and their magneto-crystalline anisotropy is 2–4 orders of magnitude larger.

The non-volume-conserving magnetostriction along with non-hysteretic and linear magnetization resembling rotation in all directions cannot be explained by the conventional picture of JM or in terms of conventional ferromagnets. Ultimately, any mechanism of NJM must allow for volume changes and simultaneously explain the confounding directional independence of magnetization curves. Before explaining it, a clue to its origins is found in the micromagnetic structure in Fig. 3, which shows that the non-Joulian magnetostrictive strain, marked in blue (along [100]) or green (along [110]) in Fig. 1a, b, is associated with microscopic reorientation of rigid cells. Only a small nonlinear strain ~16–20 p.p.m. (marked in red) in Fig. 1a, b near saturation fields constitutes conventional or JM. The cellular

¹Materials Genomics Laboratory, Laboratory for Quantum Materials & Devices, Mechanical Engineering Department, and Temple Materials Institute, Temple University, Philadelphia, Pennsylvania 19122, USA. ²Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, USA.

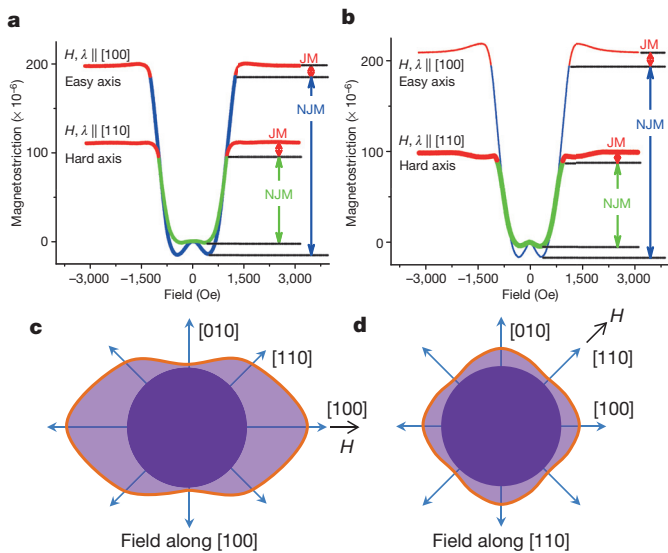


Figure 1 | Non-Joulian magnetostriction. **a**, Room-temperature longitudinal magnetostriction along magnetic easy [100] and hard [110] axes in $\text{Fe}_{73.9}\text{-Ga}_{26.1}$. The single crystal was annealed at 1,033 K for 30 min followed by rapid quenching. **b**, Room-temperature longitudinal magnetostriction along [100] and [110] axes in $\text{Fe}_{82.9}\text{-Ga}_{17.1}$, annealed at 1,033 K for 30 min then cooled at 10 K min^{-1} . The NJM (blue or green) is shown separated from JM (red). **c**, **d**, Positive volume change as a hallmark of NJM for $\text{Fe}_{73.9}\text{-Ga}_{26.1}$ alloy. The overall shapes are shown for the cases of applied field along [100] (c) and along the [110] axis (d).

magnetic structure responsible for NJM is shown in Fig. 3A. The image is notable for its exceptional magnetic periodicity (large coherence length) along the vertical [010] and the horizontal [100] directions, $\lambda_{\text{transverse}} \sim 43\text{ }\mu\text{m}$ and $\lambda_{\text{longitudinal}} \sim 12\text{--}14\text{ }\mu\text{m}$, respectively. This long-range periodicity prevails over the entire disk, which is 5 mm in diameter (Extended Data Fig. 4). A higher-magnification image, Fig. 3B, displays cells in adjacent horizontal bands with identical micromagnetic arrangements, shown schematically in Fig. 3C. However, the domain walls are mirror images of each other in the ferrofluid pattern due to their opposite chirality; domain walls with lighter contrast in a given horizontal band appear dark in adjacent bands. This reversal of spin chirality along the transverse direction reflects long-range order, as evidenced by its repetition over macroscopic distances ($\sim\text{mm}$). Note that the ripples on selected boundaries represent the signature of twins. Also notice the four-fold periodicity of layers, which is interpreted as representing the possibility and/or impossibility of observing twins end-on. We show that this structure embodies a generalized Landau structure to accommodate the large magnetostriction and differences in saturation magnetization in the base and c-directions of the underlying tetragonal state (see Supplementary Information and Extended Data Fig. 5). Occasionally, the periodicity is broken by defects resembling anti-phase boundaries (APBs) (Fig. 3D). An APB causes a precise vertical displacement of a section of the pattern relative to its adjacent regions by an amount equal to $\frac{1}{2}\lambda_{\text{transverse}}$. As a consequence, the chirality of domain walls changes along a given horizontal band. At higher magnifications, domain walls are found to be not straight but zigzagged (Fig. 3E), revealing an even finer structural heterogeneity at the sub-microscopic scale. The zigzagged walls show that the idealized schematization of magnetization in Fig. 3C in fact also follows a zigzagged pattern, as shown in a highly exaggerated schematic in Fig. 3F. A special feature of each ‘cell’ within any horizontal band, whose periodicity is labelled as $\lambda_{\text{longitudinal}}$ in Fig. 3C, is their magnetostatically and magnetoelastically autarkic character; that is, the cells represent entities that are demagnetized and de-elasticized. Here the term ‘de-elasticized’ is used to imply long-range cancellation of elastic fields analogous to the term ‘demagnetized’ in magnetism. This can be seen

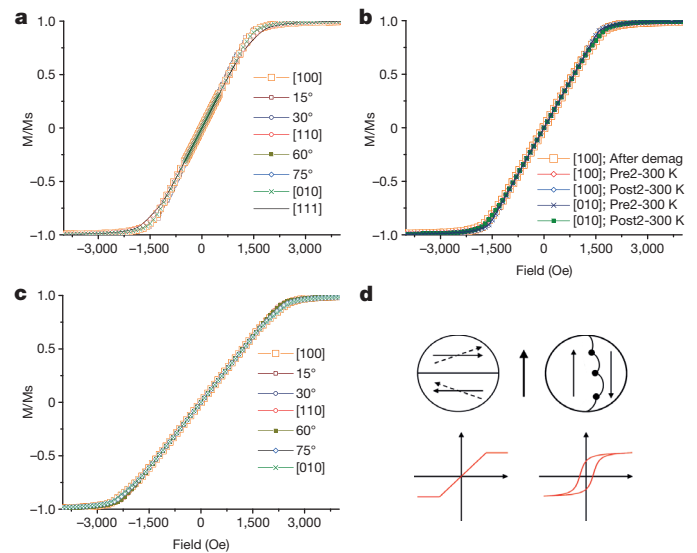


Figure 2 | Hysteresis-free, linearly reversible and isotropic magnetism. **a–c**, Room-temperature magnetization curves along various in-plane directions of disk-shaped samples. **M**, magnetization; **Ms**, saturation magnetization. **a**, Fe–Ga. The [111] direction from a bigger crystal is also shown to emphasize isotropy. **b**, **c**, Fe–Ge (b) and Fe–Al (c) alloys. **d**, Schematic showing hard axis (linear) and easy axis (hysteretic) magnetization in conventional magnets. The slopes are given by inverse demagnetization factor of the disks.

from a schematic of magnetization within each cell in Fig. 3C; the pattern notably differs from one seen in conventional magnets, in which vertical segments defining the cell boundaries are absent. A simple analysis of magnetostrictive strains further shows that they are de-elasticized (Fig. 3G), since the net force at each node of the cell is zero (nodes are stable) at zero-field.

The panel of micrographs in Fig. 3H, a–d shows snapshots of *in situ* micromagnetic studies revealing the cell’s bidirectional stability—they either exist along [100] or [010], and no other angle in between. These cells abruptly reorient as a ‘unit’ along one of the two directions when a magnetic field is applied along [100] or [010]. During micromagnetic studies we have noticed that ‘defects’ or APBs often serve as initiation sites for cell rotation. The window in Fig. 3H, a (bottom left) highlights this observation. The abrupt nature of this reorientation was further verified by Barkhausen noise measurements (H.D.C. and M.W., unpublished observations). This reorientation is facilitated by the screening through the demagnetizing and de-elasticification fields. Starting with the observation that cells are magnetostatically and magnetoelastically self-sufficient, we note that the change of the direction of magnetization at low fields is accomplished by abrupt cell reorientation and not by coherent rotation of the magnetization, as is the case for JM. The magnetocrystalline anisotropy must therefore consist of two parts, one controlled by cell reorientation and the other controlled by coherent rotation of the magnetization near saturation fields. The magneto-‘crystalline’ anisotropy energy of coherent magnetization rotation in $\text{Fe}_{82}\text{-Ga}_{17}$ equals $K_{\text{coh}} \sim 3 \times 10^4\text{ J m}^{-3}$, when determined from magnetization curves¹⁷. However, torque magnetometry¹⁸ was able to reveal an additional uniaxial anisotropy superimposed on the expected four-fold anisotropy of the crystal (A. Lisfi and M.W., unpublished observations). The two-fold anisotropy, K_{cell} , is very small, $\sim 4 \times 10^{-2} K_{\text{coh}}$, assuring cell rotation in advance of coherent rotation. We found similarly low values in Fe–Al and Fe–Ge alloys (data not shown). The Fe–17.1at%Ga alloy in Fig. 1b shows similar micromagnetic behaviour, as does Fe–26.1at%Ga. The alloy with 26.1at%Ga is chosen for illustration because it displayed fewer surface scratches. Having found similar behaviour in the two alloys on either side of the composition at which peak magnetostriction occurs ($\sim 19\text{--}$

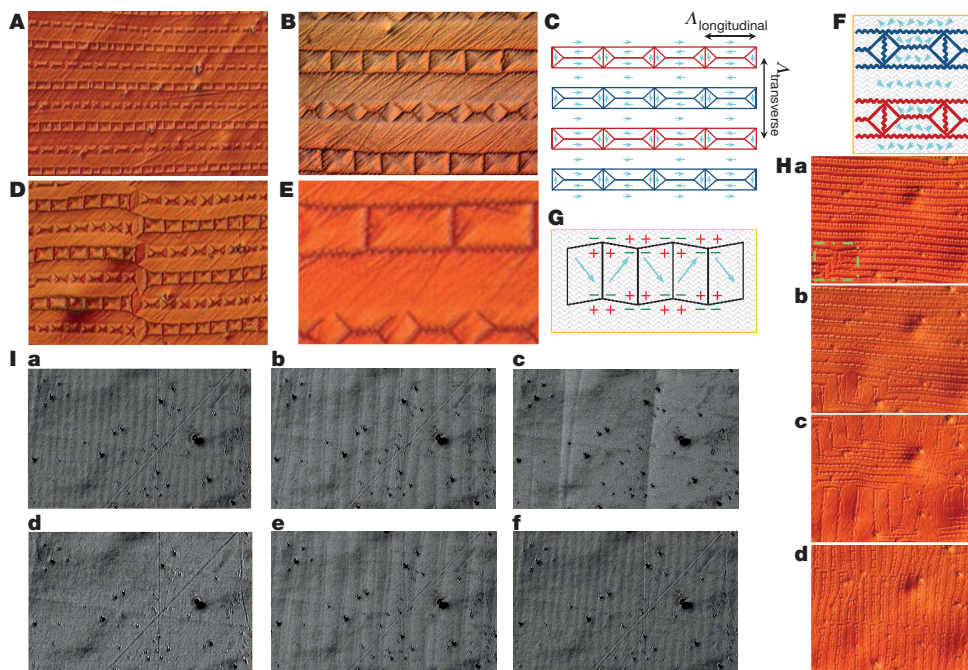


Figure 3 | Self-strain associated with highly periodic cellular micromagnetic structure gives rise to NJM. A, B, Low- and medium-magnification micromagnetic images in quenched $\text{Fe}_{73.9}\text{-Ga}_{26.1}$ crystal. The highly periodic longitudinal and transverse cellular pattern of $43\ \mu\text{m}$ and $\sim 12\text{--}14\ \mu\text{m}$, respectively, obviates the need for scale bars. Original magnifications are $\times 10$ and $\times 50$, respectively. C, Schematic of micromagnetic structure. D, Formation of APB caused by vertical displacement of periodic structure by $\frac{1}{2}A_{\text{transverse}}$. Original magnification is $\times 10$. E, Zoomed-in high-magnification image reveals that walls are zigzagged and not straight segments. Original magnification is

$\times 50$. F, Corresponding schematic in which the direction of zigzagged magnetization vectors is highly exaggerated. G, Analysis based on magnetostriction mismatch shows that average stress at nodes is zero. H, a–d, Field dependence of magnetic domains. Applied field is in horizontal, $[100]$, direction. The respective fields are 0, 186, 248, 0 Oe. Original magnification is $\times 10$. I, a–f, Evolution of microstructure in externally applied field. The field is normal to the bands. I, a–c, corresponds to increasing and I, d–f, to decreasing fields. Original magnification is $\times 5$. Respective fields are 0, 206, 393, $\sim 1,000$, 198, and 0 Oe.

20at%Ga)^{19,20}, we expect similar behaviour under suitable thermal history.

In carefully prepared Fe–Ga single crystal samples that are free of any polishing-induced deformation layers, the existence of a microscopic-scale structural heterogeneity in the form of plane-parallel bands along the $\langle 100 \rangle$ -type crystal direction becomes apparent, as shown in the differential interference contrast (DIC) optical micrograph in Fig. 3I, a–f. The images were recorded on a ‘free’ sample, that is, without constraining the sample with an

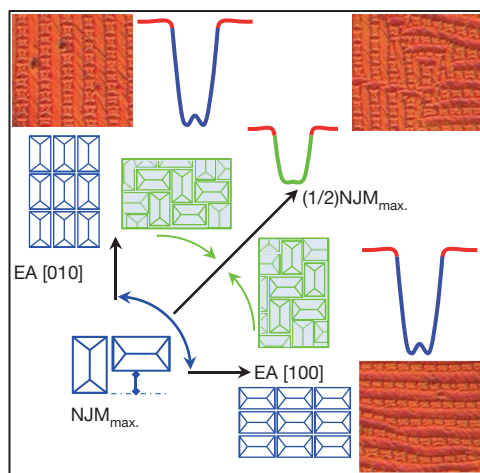


Figure 4 | Rule of mixture explaining the angular dependence of NJM. The rule is based on the fact that cells either exist along the $[100]$ or $[010]$ directions. EA, easy axis.

underlying substrate. In other words, the material spontaneously exists in this folded state at zero-field. The panel of micrographs in Fig. 3I, a–f shows that the folded structure can be progressively ‘unfolded’ by applying magnetic fields of increasing strength (Fig. 3I, a–c). The applied field is directed horizontally, approximately perpendicular to the bands. When the field strength is reduced back to zero, the crystal spontaneously returns to its folded state (Fig. 3I, d–f), that is, the cellular structure represents an equilibrium state. This unfolding and folding process occurs by motion of bands normal to themselves and their coalescence. After reducing the field strength to zero (Fig. 3I, f), the periodicity and orientation of the bands remains unchanged relative to Fig. 3I, a, indicating that they are of crystallographic origin. As-quenched stresses could be present. However, we believe that the extraordinarily long coherence length of the micromagnetic structure (Fig. 3a and Extended Data Fig. 4) indicates that they do not interfere with the formation of the pattern, its reproducibility (Extended Data Fig. 6), or the observation of NJM. In addition, in one of the experiments, samples were slowly heated to $\sim 500\ \text{K}$ ($\sim 1\ \text{K min}^{-1}$) to eliminate potential quenching stresses. The micropatterns taken before and after this experiment remain unchanged (H.D.C., unpublished observations). The observed micromagnetic images are highly reproducible, that is, they represent an equilibrium state (Extended Data Fig. 6). The DIC images (a versus f in Fig. 3I) show similar reproducibility after cycling in applied fields.

The magnitude of observed NJM in these crystals, including the observation that $\lambda_{110} \approx \frac{1}{2}\lambda_{100}$ and is always of positive sign can now be explained on the basis of a simple rule of mixture based on the population of cells along the two principal axes, $[100]$ and $[010]$, as summarized in Fig. 4. Initially, consider all the cells to be aligned along the $[100]$ axis. When $H \parallel [110]$, cells do not rotate along $[110]$. Instead,

half the cells reorient along [010], the other half remain along [100], and a strain of 100 p.p.m. is realized along [110] equal to half the strain along the principal [100] axis. Any further increase in field beyond this point causes destruction of cells due to coherent rotation associated with small JM. When $H \parallel [010]$, all the cells reorient along [010], and a maximum strain of 200 p.p.m. associated with the self-strain of the cells is observed. It is followed by a small JM associated with destruction of cell structure at high fields.

Any volume increase must necessitate changes of the lattice constants of the crystal. The conventional adaptive mechanisms such as JM or self-accommodation of ferroelastic domains in the martensite state of shape memory alloys⁵ are all volume conserving. In NJM alloys, the highly periodic magnetoelastic structure over macroscopic distances is in fact composed of microscopic repeat cells whose boundaries are defined by zigzagged domain walls of sub-microscopic segments. The zigzagged walls demonstrate that magnetization within the cells is not uniform (Fig. 3F). Stated differently, the lattice is modulated. Such a modulation involves strain gradients that relax in an applied field, giving rise to change in the volume of the crystal. It has been shown that a term $\mathbf{f} \cdot \nabla(\boldsymbol{\varepsilon} \times \boldsymbol{\eta})$ (ref. 21), where $\boldsymbol{\varepsilon}$ represents strain and $\boldsymbol{\eta}$ an order parameter, such as a spontaneous polarization, accounts for the macro- and nano-flexostriction in ferroelectrics²². We propose that the amplitude of charge density waves (CDWs)^{11,12} represents the order parameter that can describe the atomic origin of the macroscopically modulated state we have discovered; the direct replacement of the polarization order parameter by the magnetization would violate time inversion symmetry. The long-range periodicity is a hallmark of CDWs, which are also well known to accompany a soft TA_2 phonon mode as observed in Fe–Ga²³. Briefly, a CDW is a modulation of the conduction electrons that occurs concurrent with the lattice distortion. In such systems the increase in elastic energy is offset by a greater decrease in energy of the conduction electrons. The coupled distortion of the conduction electron density and the crystal leads to an overall decrease in the energy of the crystal, and equilibrium is reached when $\frac{\partial}{\partial A}(E_{\text{elec.}} + E_{\text{elas.}}) = 0$; $E_{\text{elec.}}$ and $E_{\text{elas.}}$ are the electronic and elastic energy terms, respectively, and the extent of distortion A being determined by the size of the electronic energy gap. In a non-magnetic system, there is no possibility of changing the lattice modulations caused by CDW through an applied magnetic field. However, in a ferromagnetic medium, the spin will automatically redistribute. Thus, the system becomes susceptible to external magnetic fields and therefore tunable. This allows the modulation of a CDW to be unfolded continuously in a varying field, whose macroscopic manifestation can be seen in Fig. 3I, a–f. Other mechanisms may give rise to lattice modulation. Those leading to a volume increase include volume magnetostriction, and forced magnetostriction associated with a strain gradient of exchange energy or the magnetostatic form factor. However, these mechanisms yield only miniscule volume increases², much smaller than those observed here.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 January; accepted 27 March 2015.

1. Joule, J. P. On the effects of magnetism upon the dimensions of iron and steel bars. *Philos. Mag.* **30**, 76–87 (1847).
2. du Trémolet de Lacheisserie, É., Gignoux, D. & Schlenker, M. *Magnetism: Fundamentals, Materials and Applications* (Springer, 2002).

3. Khachatryan, A. & Shatalov, G. Theory of macroscopic periodicity for a phase transition in the solid state. *Sov. Phys. JETP* **29**, 557–561 (1969).
4. Clark, A. E. & Belson, H. S. Giant room-temperature magnetostrictions in TbFe_2 and DyFe_2 . *Phys. Rev. B* **5**, 3642 (1972).
5. Chopra, H. D., Bailly, C. & Wuttig, M. Domain structures in bent In-22.5 at.%Ti polydomain crystals. *Acta Mater.* **44**, 747–751 (1996).
6. Roytburd, A. L. Thermodynamics of polydomain heterostructures. I. Effect of macrostresses. *J. Appl. Phys.* **83**, 228–238 (1998).
7. Chopra, H. D., Ji, C. H. & Kokorin, V. V. Magnetic-field-induced twin boundary motion in magnetic shape-memory alloys. *Phys. Rev. B* **61**, R14913–R14915 (2000).
8. Sozinov, A., Likhachev, A. A., Lanska, N. & Ullakko, K. Giant magnetic-field-induced strain in NiMnGa seven-layered martensitic phase. *Appl. Phys. Lett.* **80**, 1746–1748 (2002).
9. Jin, Y. M., Wang, Y. U., Khachatryan, A. G., Li, J. F. & Viehland, D. Adaptive ferroelectric states in systems with low domain wall energy: Tetragonal microdomains. *J. Appl. Phys.* **94**, 3629–3640 (2003).
10. Kaufmann, S. *et al.* Adaptive modulations of martensites. *Phys. Rev. Lett.* **104**, 145702 (2010).
11. Grüner, G. *Density Waves in Solids* (Perseus, 2000).
12. Grüner, G. The dynamics of charge-density waves. *Rev. Mod. Phys.* **60**, 1129–1181 (1988).
13. Chopra, H. D. & Wuttig, M. Non-Joulian magnetostriction (poster and abstract). *Proc. MSE Conf. Symp. A03* (23–25 September, 2014).
14. Guruswamy, S., Srisukhumbowornchai, N., Clark, A. E., Restorff, J. B. & Wun-Fogle, M. Strong, ductile, and low-field-magnetostrictive alloys based on Fe–Ga. *Scr. Mater.* **43**, 239–244 (2000).
15. Clark, A. E., Restorff, J. B., Wun-Fogle, M., Lograsso, T. A. & Schlager, D. L. Magnetostrictive properties of body-centered cubic Fe–Ga and Fe–Ga–Al alloys. *IEEE Trans. Magn.* **36**, 3238–3240 (2000).
16. Sullivan, M. R. & Chopra, H. D. Temperature- and field-dependent evolution of micromagnetic structure in ferromagnetic shape-memory alloys. *Phys. Rev. B* **70**, 094427 (2004).
17. Rafique, S., Cullen, J. R., Wuttig, M. & Cui, J. Magnetic anisotropy of FeGa alloys. *J. Appl. Phys.* **95**, 6939–6941 (2004).
18. Lisfi, A., Pokharel, S., Morgan, W., Warren, G. & Wuttig, M. The power of torque magnetometry: defect induced switching in hexaferrite nano-structures. *Nanotechnology* **25**, 415702 (2014).
19. Xing, Q., Du, Y., McQueeney, R. J. & Lograsso, T. A. Structural investigations of Fe–Ga alloys: phase relations and magnetostrictive behavior. *Acta Mater.* **56**, 4536–4546 (2008).
20. Clark, A. E. *et al.* Extraordinary magnetoelasticity and lattice softening in bcc Fe–Ga alloys. *J. Appl. Phys.* **93**, 8621–8623 (2003).
21. Pirc, R., Blinc, R. & Scott, J. F. Mesoscopic model of a system possessing both relaxor ferroelectric and relaxor ferromagnetic properties. *Phys. Rev. B* **79**, 214114 (2009).
22. Fu, J. Y. & Cross, L. E. On the flexoelectric effects in solid dielectrics: theories and applications. *Ferroelectrics* **354**, 238–245 (2007).
23. Wuttig, M., Dai, L. & Cullen, J. Elasticity and magnetoelasticity of Fe–Ga solid solutions. *Appl. Phys. Lett.* **80**, 1135–1137 (2002).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors appreciate the donation of single crystals by T. A. Lograsso. We thank S. Fähler for discussions. We thank S. Z. Hua for discussions, input on experimental design and help with magnetic and temperature dependent measurements. We thank J. Steiner, T. Ren, E. Gande and J. N. Armstrong for help with vibrating sample magnetometry and magnetostriction measurements, and assistance in collation of some images, experiments and sample preparation. M.W. acknowledges support from the National Science Foundation DMR–Metals grant 1206397 and H.D.C. acknowledges support from the National Science Foundation DMR–Condensed Matter Physics grants 1541236 (previously 1309712) and 0964830.

Author Contributions Both authors jointly developed the concepts of NJM and reversibly linear non-dissipative magnets. Both authors jointly designed the experiments. Both authors contributed equally to writing this manuscript, analysing and interpreting the images and data. Micromagnetic studies were done by the first author (H.D.C.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.D.C. (chopra@temple.edu) or M.W. (wuttig@umd.edu).

METHODS

Single crystals. All single crystals in this study (Fe–Ga, Fe–Ge and Fe–Al) were cut from boules grown at Ames laboratory into circular disks ~ 5 mm in diameter and 0.4–0.5-mm thick. The crystallographic [001] direction is normal to the disk plane.

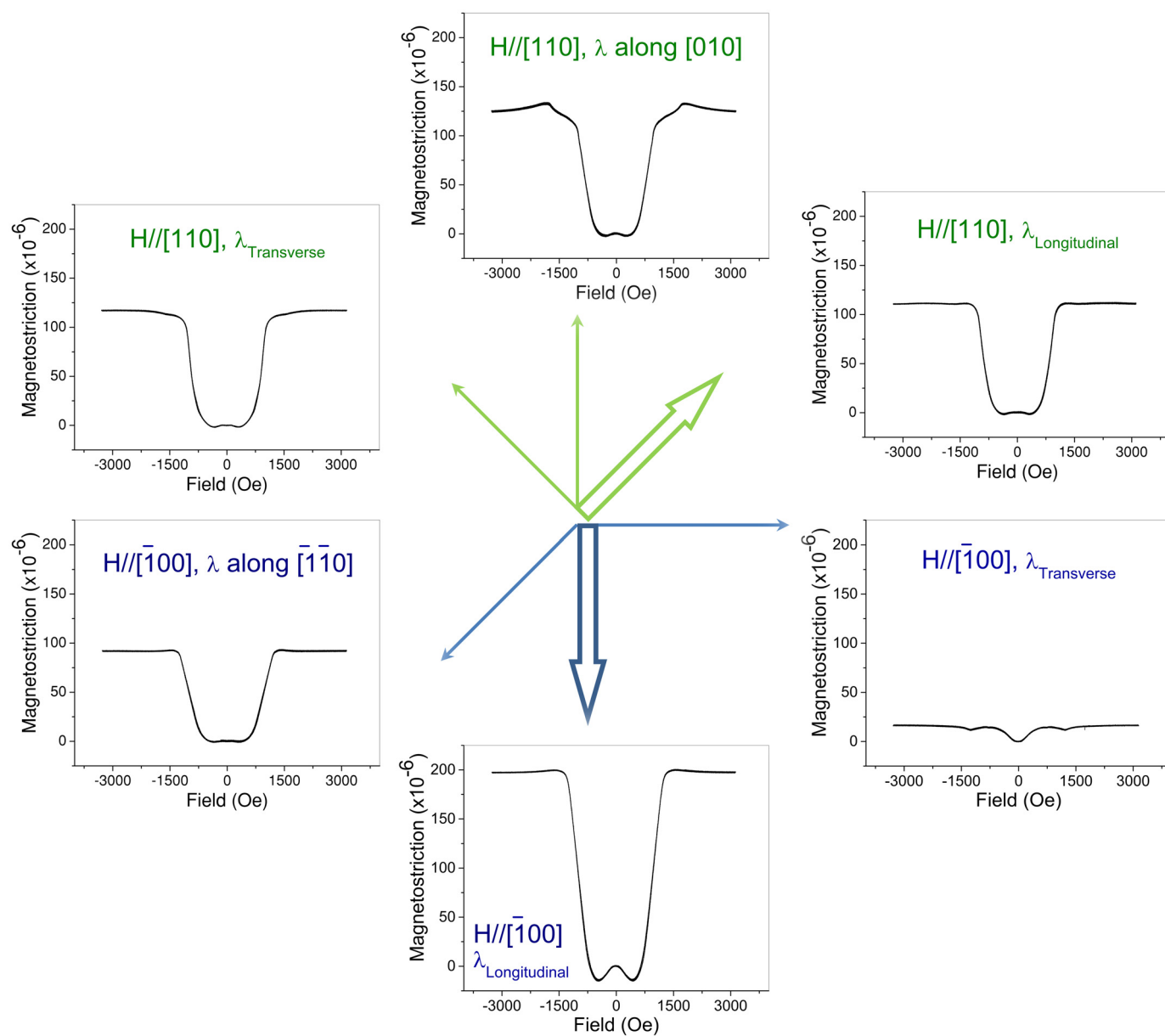
Surface preparation for micromagnetic and optical studies. It is not commonly appreciated that the lower the anisotropy (especially in our case, in which anisotropy is of purely magnetoelastic origin), the harder it is to develop the micro-magnetic pattern by surface polishing. Given the delicate nature of the patterns, their long (\sim mm) coherence length and the exceptional shear softness of the alloys, great care was taken in surface preparation and sample handling. Key precautions are as follows. Even a soft polishing cloth can lead to deep buried subsurface deformation, which upon etching (50:50 nitric acid and distilled water solution) reveals itself and destroys any subsequent pattern (compare with Extended Data Figs 7a, b). Use of tweezers to hold the samples can create long-range subsurface deformation. Samples were manipulated by placing them on soft tissues for transferring onto surfaces. Although colloidal silica (0.05 μ m) is generally recommended for polishing soft materials, it tends to agglomerate and leave deep polishing marks; similar agglomeration problems as with colloidal silica were encountered if finer alumina (0.05 μ m) was used as the last polishing step. The 0.3 μ m alumina was found to be optimum as the last polishing step. The contrast associated with surface relief of the pattern is weak (barely visible in bright-field microscopy). Hence the DIC mode of the microscope was used for observations. Additionally, image background was subtracted to enhance contrast.

Micromagnetic and optical studies. The magnetic domain structure was studied using the high-resolution ICC method, which is described in detail in several previous publications^{7,16}. Briefly, the ICC method employs a colloidal solution to decorate the microfield on the surface of a ferromagnet, similar to the versatile Bitter method. However, the technique differs in the manner in which the colloid-decorated microfield is detected. In the Bitter method, a problem in contrast

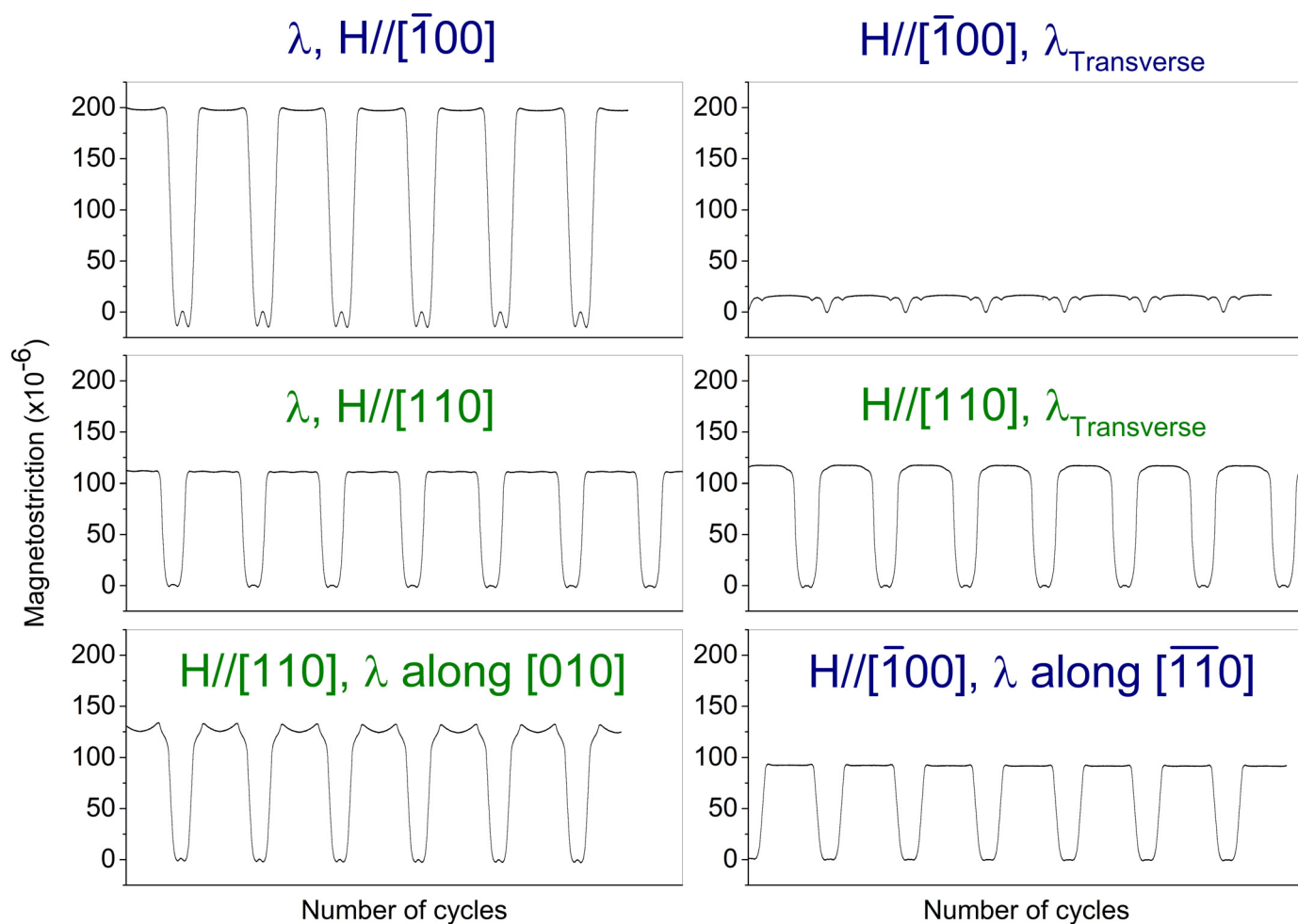
develops in the bright-field or the dark-field mode due to backscattering by particles and various surfaces between the objective lens and the specimen, resulting in an overall loss of resolution. Instead, the ICC method uses a Nomarski interferometer to detect the surface microfield distribution. The magnetic microfield on the surface causes local variation in the density of colloid particles (average colloid particle size 7 nm), thereby delineating the domain structure. This microfield is detected in the reflection mode by the interferometer optics, which detects any unevenness at the nanometer scale and reveals domain structures with a pronounced three-dimensional effect and at a high resolution that is limited only by that of the microscope (0.4–0.6 μ m). The system for imaging magnetic domains is fully automated and interfaced with an image frame grabber and a high-speed data acquisition card. Fields were measured by a Hall probe sitting underneath the sample, and tend to underreport slightly (estimated).

Magnetic measurements. Magnetostriction was measured on 5 mm Fe–Ga disks by strain gauges from Omega (KFG-2N-120-C1-11L1M2R, Linear, gauge length 2.0 mm, discontinued model) and Vishay (C2A-XX-062LT-120, Tee Rosette, gauge length 1.52 mm; WK-XX-030WT-120, Tee Rosette, gauge length 0.76 mm) using the Wheatstone bridge. Various gauge lengths were used (0.76 mm, 1.52 mm, 2 mm) to confirm the measurements^{24,25}. Tests were also made on reference nickel samples to confirm measurement protocols. The highly compliant strain gauges do not impede the evolution of magnetostriction strains. The unbalance signal of a Stanford Research Model 810 DSP amplifier provided the strain.

24. Sullivan, M. Wheatstone bridge technique for magnetostriction measurements. *Rev. Sci. Instrum.* **51**, 382–383 (1980).
25. Lisfi, A., Ren, T., Khachatryan, A. & Wuttig, M. Nano-magnetism of magnetostriction in Fe₃₅Co₆₅. *Appl. Phys. Lett.* **104**, 092401 (2014).

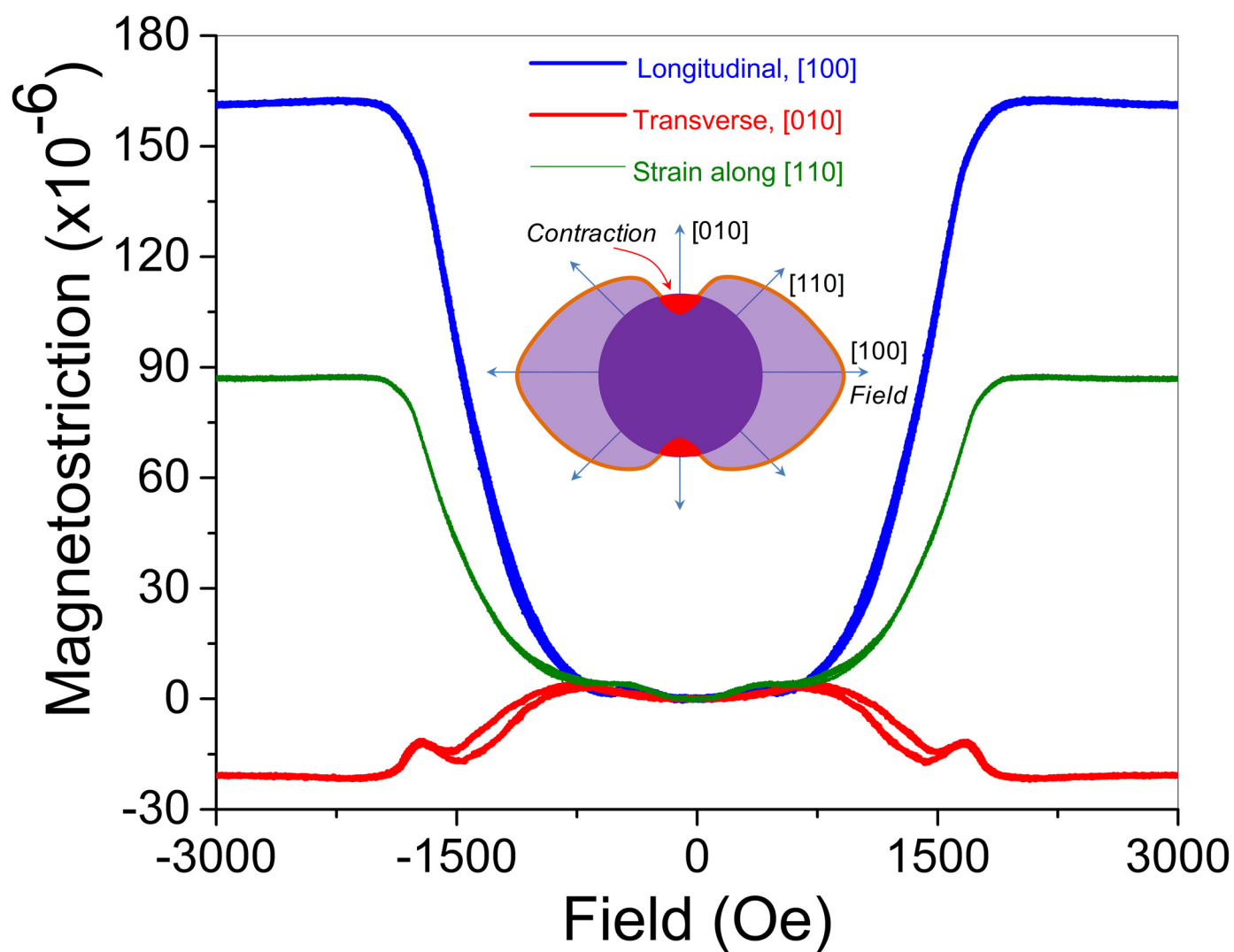


Extended Data Figure 1 | Anisotropy of magnetostriction. The data displays the angular dependence of magnetostriction along various directions in an as-quenched $\text{Fe}_{73.9}\text{Ga}_{26.1}$ single crystal.



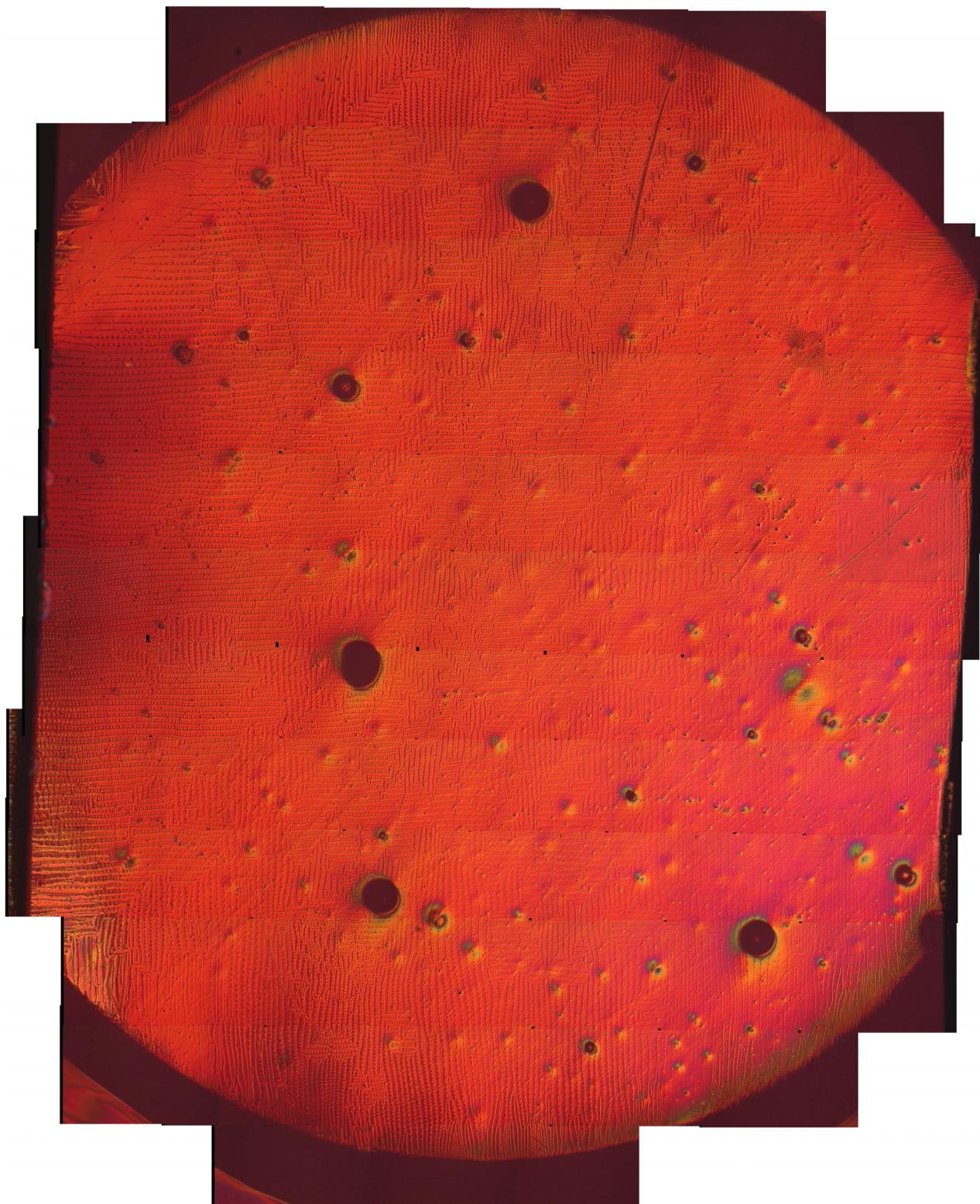
Extended Data Figure 2 | Reproducibility of magnetostriction curves. Reproducibility is shown for various traces in Fig. 1a and traces in Extended Data Fig. 1. Maximum field for each cycle is approximately $\pm 3,150$ Oe.

Similarly reproducible traces were observed for the $\text{Fe}_{82.9}\text{Ga}_{17.1}$ single crystal in Fig. 1b, not shown.



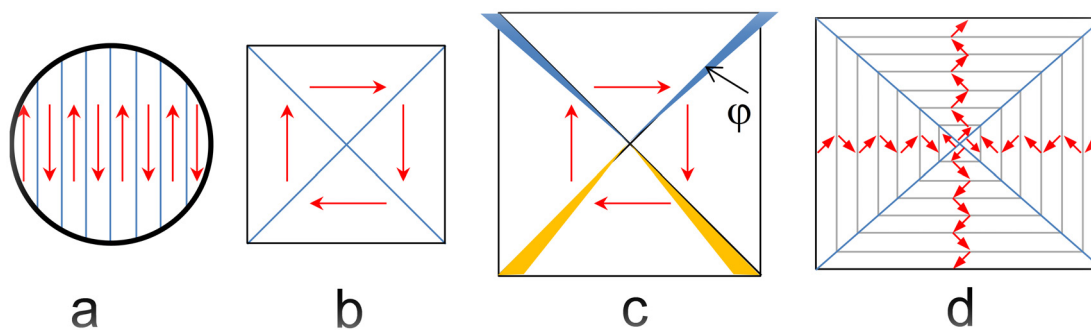
Extended Data Figure 3 | Degradation of NJM characteristics in $\text{Fe}_{73.9}\text{Ga}_{26.1}$ single crystal when cooled slowly from 1,033 K to room temperature (furnace cooled). In comparison to volume expansion in all directions when an alloy of this composition was rapidly quenched (Fig. 1a), furnace (slow)

cooling cause transverse magnetostriction to become slightly contractile. However, unlike conventional ferromagnets, the sample still exhibits a net volume increase, that is, NJM.



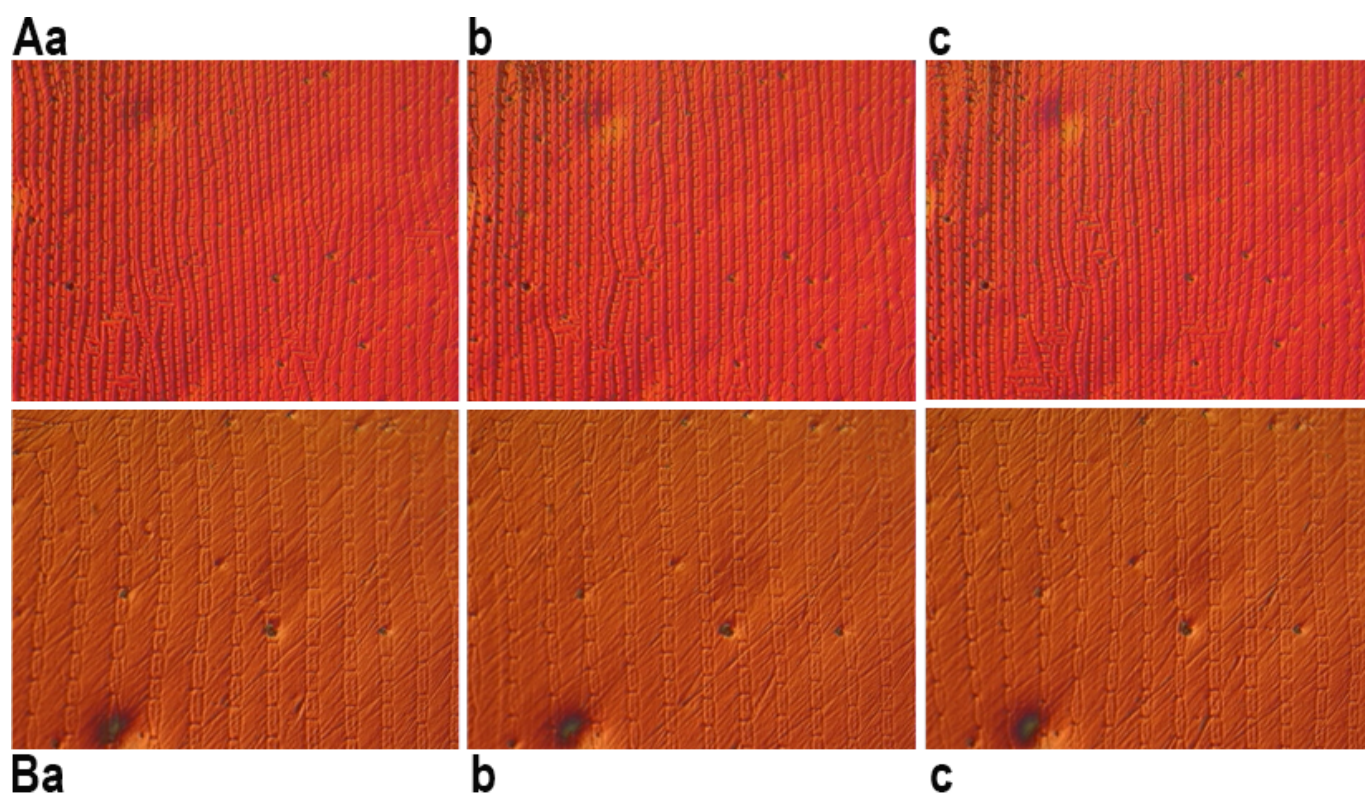
Extended Data Figure 4 | Collage showing magnetic domains across the entire 5-mm-diameter circular single crystal sample of $\text{Fe}_{73.9}\text{Ga}_{26.1}$, which was rapidly quenched from 1,033 K to room temperature. The collage was prepared after polishing and etching the sample but before applying any

magnetic field. Notice the existence of micromagnetic motifs along both [100] and [010] axes. Also notice the existence of APBs. The collage consists of high-resolution images and can thus be zoomed in for further analysis by the scientific community. Original magnification is $\times 5$.



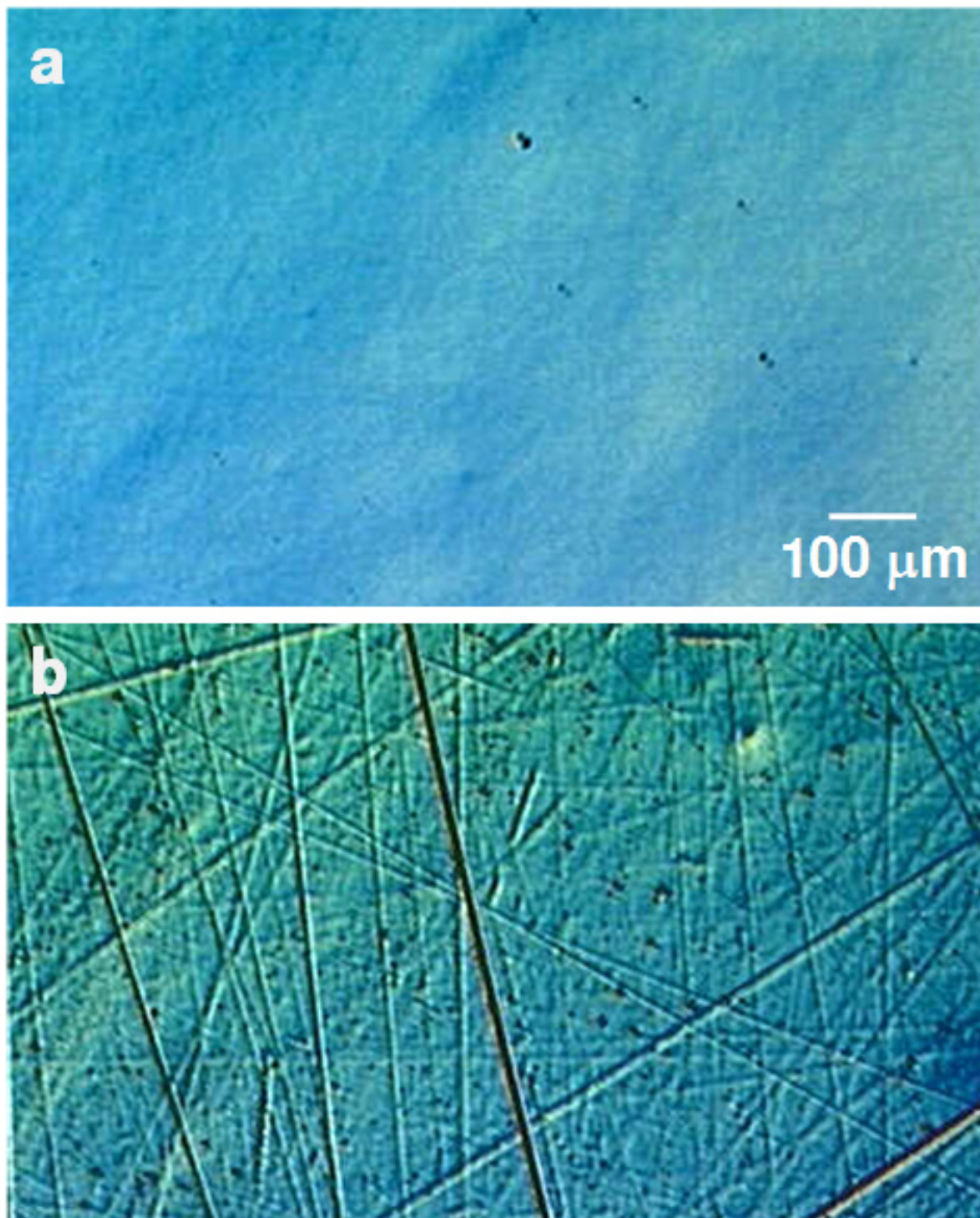
Extended Data Figure 5 | Origin of magnetically and elastically compensated state. **a**, Demagnetized state of a thin circular plate. **b**, Demagnetized state of a thin ferromagnetic film, $\lambda_{100} = 0$. **c**, Angular

distortion of the diagonals for $\lambda_{100} < 0$ and $\lambda_{100} > 0$. **d**, Stress-free rectangular demagnetized state through twinning.



Extended Data Figure 6 | Zero-field micromagnetic patterns of as-quenched $\text{Fe}_{73.9}\text{Ga}_{26.1}$ crystal after cycling in saturation magnetic field. A, a–c, Low-magnification images. Original magnification is $\times 10$.

B, a–c, High-magnification images. The pattern's periodicity and scale equals that shown in Fig. 3. Original magnification is $\times 20$.



Extended Data Figure 7 | Polishing can cause deep buried subsurface deformation. **a**, An apparently scratch-free Fe_{82.9}-Ga_{17.1} single crystal after

polishing but before etching. **b**, Subsequent etching in a 50:50 nitric acid and distilled water solution reveals numerous scratches.

Selection on noise constrains variation in a eukaryotic promoter

Brian P. H. Metzger^{1*}, David C. Yuan^{2*}, Jonathan D. Gruber¹, Fabien Duveau¹ & Patricia J. Wittkopp^{1,2}

Genetic variation segregating within a species reflects the combined activities of mutation, selection, and genetic drift. In the absence of selection, polymorphisms are expected to be a random subset of new mutations; thus, comparing the effects of polymorphisms and new mutations provides a test for selection^{1–4}. When evidence of selection exists, such comparisons can identify properties of mutations that are most likely to persist in natural populations². Here we investigate how mutation and selection have shaped variation in a *cis*-regulatory sequence controlling gene expression by empirically determining the effects of polymorphisms segregating in the *TDH3* promoter among 85 strains of *Saccharomyces cerevisiae* and comparing their effects to a distribution of mutational effects defined by 236 point mutations in the same promoter. Surprisingly, we find that selection on expression noise (that is, variability in expression among genetically identical cells⁵) appears to have had a greater impact on sequence variation in the *TDH3* promoter than selection on mean expression level. This is not necessarily because variation in expression noise impacts fitness more than variation in mean expression level, but rather because of differences in the distributions of mutational effects for these two phenotypes. This study shows how systematically examining the effects of new mutations can enrich our understanding of evolutionary mechanisms. It also provides rare empirical evidence of selection acting on expression noise.

The *TDH3* gene encodes a highly expressed enzyme involved in central glucose metabolism⁶. Deletion of this gene decreases fitness⁷ and its overexpression alters phenotypes⁸, suggesting that the promoter controlling its expression is subject to selection in the wild. To test this hypothesis, we sequenced a 678 base pair (bp) region containing the *TDH3* promoter (P_{TDH3}) as well as the 999 bp coding sequence of *TDH3* in 85 strains of *S. cerevisiae* sampled from diverse environments (Supplementary Table 1). We observed 44 polymorphisms in P_{TDH3} : 35 single nucleotide polymorphisms (SNPs) at 33 different sites and nine insertions or deletions (indels) ranging from 1 to 32 bp (Extended Data Fig. 1a). This frequency of polymorphic sites was significantly lower than the frequency of synonymous polymorphisms within the *TDH3* coding sequence ($P = 0.03$, Fisher's exact test) and polymorphic sites were less conserved between species than non-polymorphic sites in the promoter ($P = 5 \times 10^{-5}$, Wilcoxon rank sum test), consistent with purifying selection acting on P_{TDH3} . To determine whether the polymorphisms observed in P_{TDH3} contribute to *cis*-regulatory variation, we compared relative *cis*-regulatory activity between each of 48 strains and a common reference strain. We found significant differences in *cis*-regulatory activity among strains (Extended Data Fig. 1b), and 97% of the heritable *cis*-regulatory variation could be explained by sequence variation within the *TDH3* promoter (see Methods). These differences in *cis*-regulation act together with differences in *trans*-regulation to produce variation in *TDH3* messenger RNA (mRNA) abundance observed among strains (Extended Data Fig. 1b).

To quantify the effect of each individual polymorphism on *cis*-regulatory activity, we used parsimony to reconstruct the evolutionary

relationships among the 27 P_{TDH3} haplotypes observed in the 85 strains of *S. cerevisiae* sampled. We then inferred the most likely ancestral state for these haplotypes using P_{TDH3} sequences from an additional 15 strains of *S. cerevisiae* and all known species in the *Saccharomyces sensu stricto* genus (Supplementary Table 1 and Extended Data Fig. 2a). Next, we measured *cis*-regulatory activity of P_{TDH3} for the inferred ancestral state, each observed haplotype, and both possible intermediates between all pairs of observed haplotypes that differed by two mutational steps. We did this by cloning each P_{TDH3} haplotype upstream of the coding sequence for a yellow fluorescent protein (YFP), integrating these reporter genes (P_{TDH3} -YFP) into the *S. cerevisiae* genome, and quantifying YFP fluorescence using flow cytometry⁹. For each genotype, YFP fluorescence was measured in approximately 10,000 single cells from each of nine biological replicate populations (Fig. 1a). We used these data to estimate both mean expression level (μ ; Fig. 1b) and expression noise (σ/μ ; Fig. 1c) of P_{TDH3} -YFP for each promoter haplotype as readouts of *cis*-regulatory activity. We then inferred the effects of individual polymorphisms by comparing the phenotypes of ancestral and descendent haplotypes that differed by only a single sequence change.

To determine how the effects of P_{TDH3} polymorphisms compare with the effects of new mutations in this *cis*-regulatory element, we estimated the distribution of mutational effects by using site-directed mutagenesis to introduce 236 of the 241 possible G:C→A:T transitions individually into P_{TDH3} -YFP alleles and assayed their effects on *cis*-regulatory activity using flow cytometry as described above. We used G:C→A:T transitions to estimate the distribution of mutational effects because they were the most common type of SNP observed both in the *TDH3* promoter (Extended Data Fig. 1a) and genome-wide among the 85 *S. cerevisiae* strains^{10,11}. They were also the most frequent type of spontaneous point mutation observed in mutation accumulation lines of *S. cerevisiae*¹². To determine whether the effects of these mutations were likely to be representative of the effects of all types of point mutation, we analysed data from previously published studies that measured the effects of single mutations on *cis*-regulatory activity^{13–16}. We found no significant difference between the effects of G:C→A:T transitions and other types of point mutation on *cis*-regulatory activity in any of these data sets (Extended Data Fig. 3a–m). Consistent with this observation, we found no significant difference between the effects of G→A and C→T mutations on P_{TDH3} activity (mean expression level: $P = 0.73$; expression noise: $P = 0.52$; two-tailed *t*-test; Extended Data Fig. 3n, o). We also found no significant difference between the effects of G:C→A:T and other types of polymorphism (mean expression level: $P = 0.91$; expression noise: $P = 0.90$; two-tailed *t*-test; Extended Data Fig. 3p, q).

Mutations with the largest effects on mean expression level and expression noise were located within experimentally validated transcription factor binding sites (TFBS)^{17,18} (Fig. 2). All of these mutations decreased mean expression level and increased expression noise. Outside the known TFBS, 50% of the 218 mutations tested increased mean expression level and 87% increased expression noise. Despite this difference in the shape of the distributions, a negative correlation was observed between mean

¹Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA. ²Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109, USA.

*These authors contributed equally to this work.

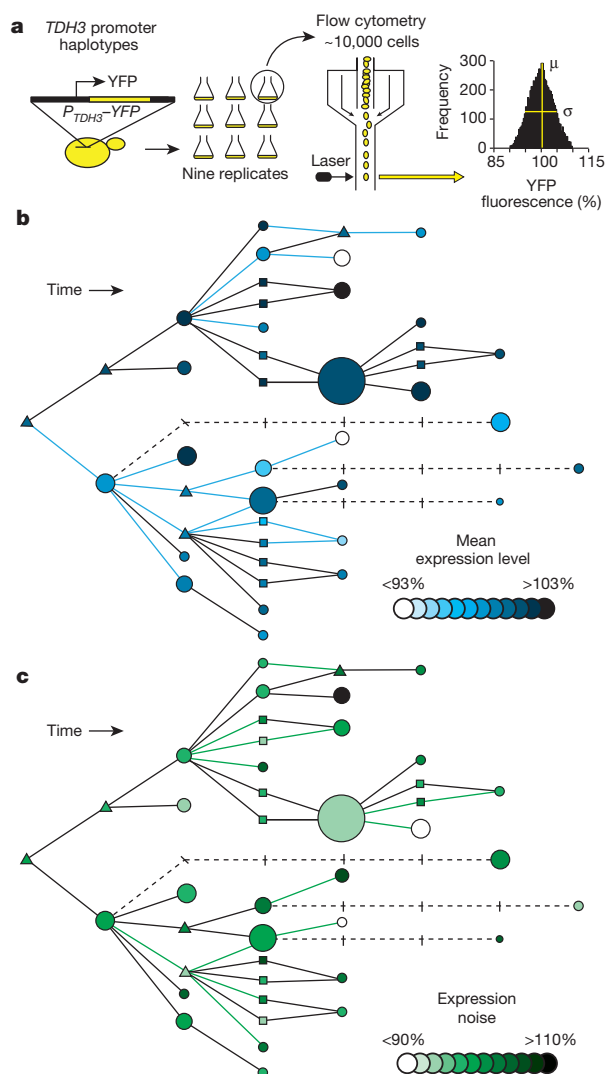


Figure 1 | Effects of polymorphisms on P_{TDH3} activity. **a**, The *cis*-regulatory activity was quantified as YFP fluorescence in nine biological replicates for each P_{TDH3} -YFP haplotype using flow cytometry. The mean (μ) and standard deviation (σ) of single-cell fluorescence phenotypes were calculated for each sample. **b**, Mean expression level of P_{TDH3} -YFP for each *TDH3* promoter haplotype is shown in the haplotype network (Extended Data Fig. 2a), with differences in mean expression level relative to the inferred common ancestor shown with different shades. Circles are haplotypes observed among the sampled strains, with the diameter of each circle proportional to frequency of that haplotype among the 85 strains. Triangles are haplotypes that were not observed among the strains sampled, but must exist, or have existed, as intermediates between observed haplotypes. Squares are possible haplotypes that might exist, or have existed, as intermediates between observed haplotypes. Dashed lines connect haplotypes by multiple mutations. On the basis of *t*-tests with a Bonferroni correction, 17 of the 45 polymorphisms present in this network caused a significant change in mean expression level (blue lines). **c**, Same as **b**, but for expression noise. Eighteen of the 45 polymorphisms present in this network caused a significant change in expression noise (green lines, *t*-test, Bonferroni corrected).

expression level and expression noise ($R^2 = 0.85$; Extended Data Fig. 4) that was similar to previous reports for other yeast promoters¹⁹. The strength of this correlation was reduced to $R^2 = 0.45$ when mutations in the known TFBS were excluded.

To take the mutational process into account when testing for evidence that selection has influenced variation in the *S. cerevisiae* *TDH3* promoter, we compared the distributions of effects for mutations and polymorphisms both of mean expression level (Fig. 3a) and of expression noise (Fig. 3b). We did this by randomly sampling sets of variants

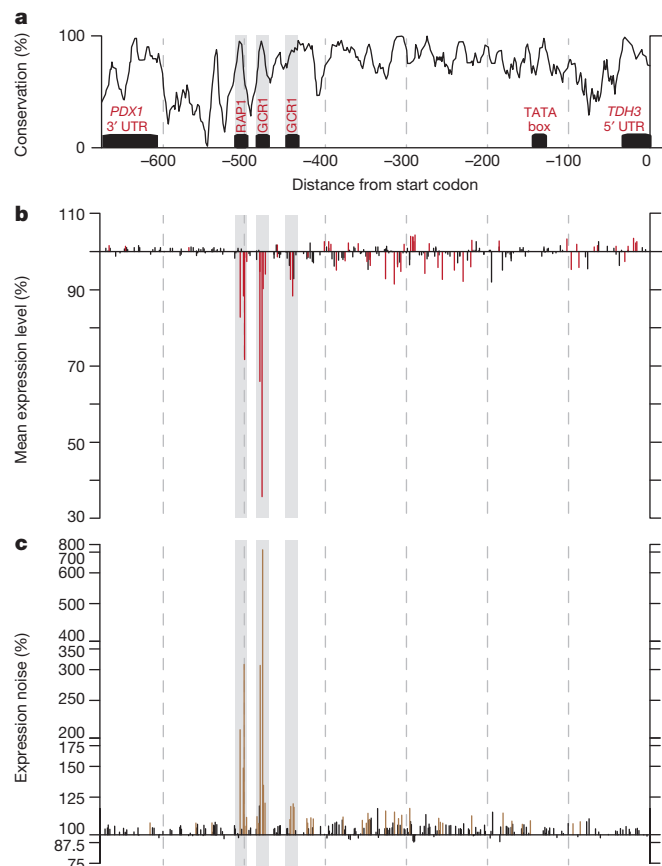


Figure 2 | Effects of mutations on P_{TDH3} activity. **a**, The structure of the 678 bp region analysed, including the *TDH3* promoter with previously identified TFBS for RAP1 and GCR1, a TATA box, and untranslated regions (UTRs) for *TDH3* and *PDX1*, is shown. The black line indicates sequence conservation across the *sensu stricto* genus. **b**, Effects of individual mutations on mean expression level are shown in terms of the percentage change relative to the un-mutagenized reference allele, and are plotted according to the site mutated in the 678 bp region. Fifty-nine of the 236 mutations tested significantly altered mean expression levels (red lines, *t*-test, Bonferroni corrected). The shaded regions correspond to the known binding sites indicated in **a**. **c**, Same as **b**, but for expression noise. Because the effects of mutations on expression noise relative to the reference allele were much greater in magnitude than the effects of these mutations on mean expression level, they are plotted on a log₂ scale. Measurements of expression noise were more variable among replicates than measurements of mean expression level, resulting in lower power to detect small changes as significant. Nonetheless, 42 of the 236 mutations tested significantly altered expression noise (brown lines, *t*-test, Bonferroni corrected).

from the mutational distribution and comparing their effects with those observed among the naturally occurring polymorphisms. We found that the effects of observed polymorphisms on mean expression level were consistent with random samples of mutations from the distribution of mutational effects (one-sided $P = 0.89$; Extended Data Fig. 5a, i), whereas the effects of observed polymorphisms on expression noise were not (one-sided $P = 0.0092$; Extended Data Fig. 5b). Specifically, polymorphisms were less likely to increase expression noise than random mutations (Extended Data Fig. 5j), suggesting that selection has preferentially retained mutations that minimize expression noise from P_{TDH3} in natural populations. These results were robust to the exclusion of the large effect mutations in known TFBS from the distribution of mutational effects and the restriction of polymorphisms to G:C→A:T changes (Extended Data Fig. 5c–f, k–n), the metric used to quantify expression noise (Extended Data Fig. 6), and differences in genetic background that include a change in ploidy from haploid to diploid (Extended Data Fig. 7).

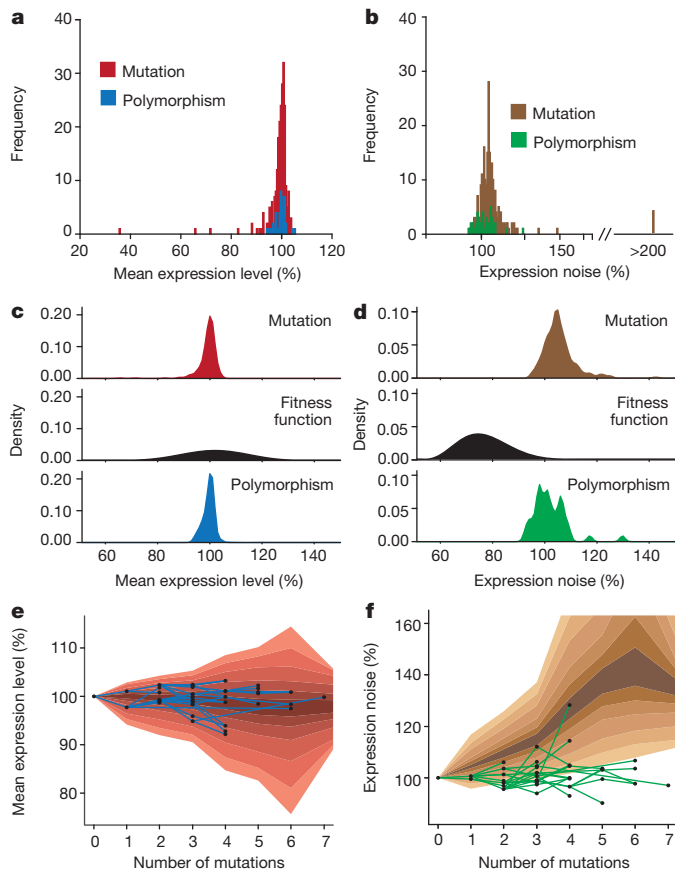


Figure 3 | Effects of selection on P_{TDH3} activity. **a**, Summary of the effects of mutations (red) and polymorphisms (blue) on mean expression level. **b**, Summary of the effects of mutations (brown) and polymorphisms (green) on expression noise. **c**, The maximum likelihood fitness function (middle, black) relating the distribution of mutational effects (top, red) to the distribution of observed polymorphisms (bottom, blue) is shown for mean expression level. **d**, Same as **c**, but for expression noise. **e**, Changes in mean expression level observed among haplotypes over time in the inferred haplotype network (Extended Data Fig. 2a) are shown in blue. The red background represents the 95th, 90th, 80th, 70th, 60th, and 50th percentiles, from light to dark, for mean expression level resulting from 10,000 independent simulations of phenotypic trajectories in the absence of selection. **f**, Same as **e**, but for expression noise. Effects of the mutational distribution are shown in brown. Expression noise among haplotypes is shown in green.

The probability that a new mutation with a particular phenotypic effect survives within a species to be sampled as a polymorphism is related to its effect on relative fitness. The function describing relative fitness for different phenotypes can therefore be inferred by comparing the distribution of effects for new mutations to the distribution of effects for polymorphisms (Fig. 3c, d). For mean expression level, we found that the most likely fitness function (Fig. 3c) did not explain the data significantly better than a uniform fitness function representing neutral evolution ($P = 0.87$). For expression noise, we rejected a model of neutral evolution ($P = 0.00019$) and found that the most likely fitness function included higher fitness for variants that decreased gene expression noise (Fig. 3d). Repeating this analysis using alternative metrics for expression noise produced comparable results (Extended Data Fig. 6). These data suggest an evolutionary model in which purifying selection preferentially removes variants that increase expression noise, resulting in robust expression of *TDH3* among genetically identical individuals.

Consistent with this model, polymorphisms with the largest effects on expression noise (but not mean expression level) were found at the lowest frequencies within the sampled strains of *S. cerevisiae* (mean, $P = 0.43$; noise, $P = 0.0029$; permutation test; Extended Data Fig. 2b, c).

However, this pattern could also result from population structure among the sampled strains. To separate the effects of selection and population structure, we used the structure of the inferred haplotype network and the distribution of mutational effects to simulate neutral trajectories for *cis*-regulatory phenotypes as they diverged from the P_{TDH3} ancestral state. We then compared these trajectories with the phenotypic changes observed among naturally occurring haplotypes and their inferred intermediates both for mean expression level (Fig. 3e) and for expression noise (Fig. 3f). We found that the observed haplotypes were consistent with neutral expectations for mean expression level (one-sided $P = 0.32$; Extended Data Fig. 5g), but were not consistent with this neutral model for expression noise (one-sided $P < 0.0001$; Extended Data Fig. 5h), regardless of which metric was used to measure expression noise (Extended Data Fig. 6). We again saw that naturally occurring haplotypes showed smaller changes in noise relative to the common ancestor than would be expected from the mutational process alone, implying persistent selection for low noise in P_{TDH3} activity in the wild.

Taken together, our data indicate that sequence variation in the *S. cerevisiae* *TDH3* promoter has been affected more by selection for low levels of noise than by selection for a particular level of *cis*-regulatory activity. This is not because the mean level of *cis*-regulatory activity is less important than noise for fitness, but because of differences in the distributions of mutational effects for these two phenotypes. Indeed, theoretical work shows that selection for low levels of noise is most likely to occur for phenotypes that are subject to purifying selection²⁰. Additional evidence suggesting that selection can act on expression noise comes from genomic analyses^{20–25} and from the conservation of ‘shadow enhancers’ that appear to maintain robust expression in multicellular organisms^{26,27}. By investigating not only the survival of the fittest, but also the ‘arrival of the fittest’^{28,29}, our work shows how phenotypic diversity produced by the mutational process itself has inherent biases that can influence the course of regulatory evolution. By taking empirical measurements of these mutational biases into account, we have identified an unexpected target of selection that impacts how a *cis*-regulatory element evolves.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 May 2014; accepted 19 January 2015.

Published online 16 March 2015.

- Smith, J. D., McManus, K. F. & Fraser, H. B. A novel test for selection on *cis*-regulatory elements reveals positive and negative selection acting on mammalian transcriptional enhancers. *Mol. Biol. Evol.* **30**, 2509–2518 (2013).
- Denver, D. R. et al. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nature Genet.* **37**, 544–548 (2005).
- Stoltzfus, A. & Yampolsky, L. Y. Climbing mount probable: mutation as a cause of nonrandomness in evolution. *J. Hered.* **100**, 637–647 (2009).
- Rice, D. P. D. & Townsend, J. P. J. A test for selection employing quantitative trait locus and mutation accumulation data. *Genetics* **190**, 1533–1545 (2012).
- Raser, J. M. & O’Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
- McAlister, L. & Holland, M. J. Differential expression of the three yeast glyceraldehyde-3-phosphate dehydrogenase genes. *J. Biol. Chem.* **260**, 15019–15027 (1985).
- Pierce, S. E., Davis, R. W., Nislow, C. & Giaever, G. Genome-wide analysis of barcoded *Saccharomyces cerevisiae* gene-deletion mutants in pooled cultures. *Nature Protocols* **2**, 2958–2974 (2007).
- Ringel, A. E. et al. Yeast Tdh3 (glyceraldehyde-3-phosphate dehydrogenase) is a Sir2-interacting factor that regulates transcriptional silencing and rDNA recombination. *PLoS Genet.* **9**, e1003871 (2013).
- Gruber, J. D., Vogel, K., Kalay, G. & Wittkopp, P. J. Contrasting properties of gene-specific regulatory, coding, and copy number mutations in *Saccharomyces cerevisiae*: frequency, effects and dominance. *PLoS Genet.* **8**, e1002497 (2012).
- Liti, G. et al. Population genomics of domestic and wild yeasts. *Nature* **458**, 337–341 (2009).
- Schacherer, J., Shapiro, J. A., Ruderfer, D. M. & Kruglyak, L. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458**, 342–345 (2009).
- Lynch, M. et al. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl Acad. Sci. USA* **105**, 9272–9277 (2008).
- Patwardhan, R. P. et al. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nature Biotechnol.* **27**, 1173–1175 (2009).

14. Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nature Biotechnol.* **30**, 271–279 (2012).
15. Patwardhan, R. P. *et al.* Massively parallel functional dissection of mammalian enhancers *in vivo*. *Nature Biotechnol.* **30**, 265–270 (2012).
16. Kwasniewski, J. & Mogno, I. Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl Acad. Sci. USA* **109**, 19498–19503 (2012).
17. Yagi, S., Yagi, K., Fukuoka, J. & Suzuki, M. The UAS of the yeast GAPDH promoter consists of multiple general functional elements including RAP1 and GRF2 binding sites. *J. Vet. Med. Sci.* **56**, 235–244 (1994).
18. Baker, H. V. *et al.* Characterization of the DNA-binding activity of GCR1: *in vivo* evidence for two GCR1-binding sites in the upstream activating sequence of *TP1* of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**, 2690–2700 (1992).
19. Hornung, G. *et al.* Noise-mean relationship in mutated promoters. *Genome Res.* **22**, 2409–2417 (2012).
20. Lehner, B. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.* **4**, 1–6 (2008).
21. Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J. & Eisen, M. B. Noise minimization in eukaryotic gene expression. *PLoS Biol.* **2**, 834–838 (2004).
22. Wang, Z. & Zhang, J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc. Natl Acad. Sci. USA* **108**, E67–E76 (2011).
23. Newman, J. R. S. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846 (2006).
24. Batada, N. & Hurst, L. Evolution of chromosome organization driven by selection for reduced gene expression noise. *Nature Genet.* **39**, 945–949 (2007).
25. Zhang, Z., Qian, W. & Zhang, J. Positive selection for elevated gene expression noise in yeast. *Mol. Syst. Biol.* **5**, 1–12 (2009).
26. Frankel, N. *et al.* Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 1–5 (2010).
27. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20**, 1562–1567 (2010).
28. Fontana, W. & Buss, L. “The arrival of the fittest”: toward a theory of biological organization. *Bull. Math. Biol.* **56**, 1–64 (1994).
29. De Vries, H. *Species and Varieties, Their Origin by Mutation* 825–826 (Open Court, 1905).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Maclean, J. Zhang, and C. Hittinger for strains, the University of Michigan Center for Chemical Genomics for technical assistance with flow cytometry, and J. Coolon, R. Lusk, K. Stevenson, A. Hodgins-Davis, J. Lachowiec, C. Maclean, J. Yang, C. Landry, J. Townsend, and D. Petrov for comments on the manuscript. Funding for this work was provided to P.J.W. by the March of Dimes (5-FY07-181), Alfred P. Sloan Research Foundation, National Science Foundation (MCB-1021398), National Institutes of Health (1 R01 GM108826) and the University of Michigan. Additional support was provided by the University of Michigan Rackham Graduate School, Ecology and Evolutionary Biology Department, and the National Institutes of Health Genome Sciences training grant (T32 HG000040) to B.P.H.M.; by a National Institutes of Health Genetics training grant (T32 GM007544) to D.C.Y.; by a National Institutes of Health National Research Service Award (NRSA) postdoctoral fellowship (1 F32 GM083513-0) to J.D.G.; and by a European Molecular Biology Organization postdoctoral fellowship (EMBO ALTF 1114-2012) to F.D.

Author Contributions D.C.Y., P.J.W., and J.G. designed the mutational spectrum project. D.C.Y. created all *P_{TDH3}-YFP* mutant strains. D.C.Y., B.P.H.M., and F.D. collected flow cytometry data. B.P.H.M. and P.J.W. designed the natural variation project. D.C.Y. created all strains with natural haplotypes and B.P.H.M. performed all other experiments. B.P.H.M. analysed all data. B.P.H.M., D.C.Y., and P.J.W. wrote the manuscript.

Author Information Flow cytometry data have been deposited in the FlowRepository under Repository ID FR-FCM-ZZBN. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.J.W. (wittkopp@umich.edu).

METHODS

Characterizing variation segregating in the *TDH3* promoter. Variation in the *TDH3* gene was determined for 85 natural isolates of *S. cerevisiae*^{10,11} (Supplementary Table 1). Sequences were obtained from each strain by PCR and Sanger sequencing using DNA extracted from diploid cells. Strains heterozygous for the *TDH3* promoter were grown on GNA plates for 12 h (5% dextrose, 3% Difco nutrient broth, 1% Oxoid yeast extract, 2% agar) and sporulated on potassium acetate plates (1% potassium acetate, 0.1% Oxoid yeast extract, 0.05% dextrose, 2% agar). Individual spores were isolated by tetrad dissection and haploid derivatives were sequenced to determine empirically the phase of the two *TDH3* promoter haplotypes. All reagents for growth of yeast cultures were purchased from Fisher unless otherwise noted. In all, the 678 bp promoter contained SNPs at 33 sites and the 238 synonymous sites contained 22 SNPs. Five non-synonymous changes were also observed among these 85 strains.

Inferring the ancestral sequence and constructing the haplotype network for *P_{TDH3}*. Promoter haplotypes (Supplementary Table 1 and Extended Data Fig. 2a) were initially aligned using Pro-Coffee³⁰, followed by re-alignment with PRANK³¹ and manual adjustment around repetitive elements and indels (Supplementary File 1). The *TDH3* promoter sequences from all *Saccharomyces sensu stricto* species^{10,32–34}, as well as an additional 15 strains of *S. cerevisiae* known to be an outgroup to the 85 focal strains³⁵, were also determined by Sanger sequencing. These sequences were used to infer the ancestral state of the *TDH3* promoter for the 85 strains with both parsimony and maximum likelihood methods implemented in MEGA 6 (ref. 36); both methods gave identical results. TCS 2.1 (ref. 37) was used to build a haplotype network for the *TDH3* promoter, with changes polarized on the basis of the inferred ancestral state (Extended Data Fig. 2a). One haplotype (HH in Supplementary Table 1) could not be confidently placed within the network and was excluded from our analysis. Sequence conservation for individual sites was determined using sequences from all seven *Saccharomyces sensu stricto* species using ConSurf³⁸ and the phylogeny from a prior study³⁹. To reduce heterogeneity in plotting, conservation was averaged over a 20 bp sliding window.

Measuring variation in *TDH3* mRNA levels and *cis*-regulatory activity. *TDH3* mRNA levels and *cis*-regulatory activity were measured using pyrosequencing, with relative allelic expression in F₁ hybrids providing a readout of relative *cis*-regulatory activity⁴⁰. This technique requires one or more sequence differences to compare relative genomic DNA (gDNA) or complementary DNA (cDNA) abundance between two strains or two alleles within the same strain⁴¹. We therefore constructed reference strains of both mating types that carried a copy of the *TDH3* gene with a single, synonymous mutation (T243G). These genotypes were constructed by inserting the *URA3* gene into the native *TDH3* coding region in strains BY4741 and BY4742 and then replacing *URA3* with the modified *TDH3* coding sequence using the lithium acetate method and selection on 5-FOA^{9,42}. To do this, 80 bp oligonucleotides, containing a synonymous mutation and homology to each side of the target site, were transformed into these strains. Successful transformants (strains YPW342 and YPW339, respectively) were confirmed by Sanger sequencing. Resistance markers for hygromycin B (*hphMX6*) and G418 (*kanMX4*) were then inserted into the *HO* locus of these strains (producing YPW360 and YPW361, respectively) and used to construct a diploid reference strain (YPW362). A *kanMX4* resistance marker was also successfully inserted into the *HO* locus of 63 of the 85 natural strains^{10,11}.

To construct hybrids suitable for measuring *cis*-regulatory activity of natural isolates relative to a reference strain, haploid cells from each of the 63 natural isolates with a *kanMX4* resistance marker (mating type α) were mixed with an equal number of haploid cells from the reference strain YPW360 (mating type α) on YPD plates (2% dextrose, 1% Oxoid yeast extract, 2% Oxoid peptone, 2% agar). After 24 h, cultures were streaked on YPD plates to obtain single colonies and then patched to YPD plates containing G418 and Hygromycin B to select for diploids. Four replicates of each hybrid were grown in 500 μ l of YPD liquid media for 20 h at 30 °C in 2 ml 96-well plates with 3 mm glass beads, shaking at 250 rpm. Cultures were diluted to an attenuation, $D_{600\text{ nm}}$, of 0.1 and then grown for an additional 4 h. Plates were centrifuged, and the YPD liquid was removed. Cultures were then placed in a dry ice/ethanol bath until frozen and stored at –80 °C. To prepare samples for measuring total *TDH3* mRNA abundance in each natural isolate relative to a common reference strain, diploids for each of the 63 natural isolates were mixed with a similar number of diploid cells from strain YPW362 on the basis of OD₆₀₀ readings after the initial growth in YPD liquid. These co-cultures were incubated and processed as described above.

For each hybrid and co-culture sample, gDNA and RNA were sequentially extracted from a single lysate using a modified protocol of Promega's SV Total RNA Isolation System. After thawing cultures on ice for about 30 min, 175 μ l of SV RNA lysis buffer (with β -mercaptoethanol), 350 μ l of double-distilled water, and 50 μ l of 400 micron RNase free beads were added to each sample. Plates were vortexed until cell pellets were completely resuspended. The plates were then centrifuged and

175 μ l of supernatant was mixed with 25 μ l of RNase-free 95% ethanol and loaded onto a binding plate. To extract RNA, 100 μ l of RNase-free 95% ethanol was added to the flow through and loaded onto a second binding plate. These plates were then washed twice with 500 μ l of SV RNA wash solution and allowed to dry. To extract DNA, the first binding plate was washed twice with 700 μ l of cold 70% ethanol and allowed to dry. For both binding plates, 100 μ l of double-distilled water was added to each well, the plate was incubated at 25 °C for 7.5 min, and the elution was collected. RNA from each sample was converted to cDNA by mixing 5 μ l of extracted RNA with 2 μ l RNase-free water, 1 μ l DNase buffer, 1 μ l RNasin Plus, and 1 μ l DNase 1 and incubating at 37 °C for 1 h followed by 65 °C for 15 min. Three microlitres of oligo dT (T₁₉VN) was added and cooled to 37 °C over 35 min. Four microlitres of First Strand Buffer, 2 μ l dNTPs, 0.5 μ l RNasin Plus, and 0.5 μ l of SuperScript II were added and incubated for 1 h. Thirty microlitres of double-distilled water was then added to each sample.

Pyrosequencing was performed as described previously⁴¹ using a PSQ 96 pyrosequencing machine and Qiagen pyroMark Gold Q96 reagents for gDNA and cDNA samples both for hybrids and for co-cultured diploids. One microlitre of cDNA or gDNA was used in each PCR reaction, with primers shown in Supplementary Table 2. A single PCR and pyrosequencing reaction was performed for each gDNA and cDNA sample from each of the four biological replicate hybrid and co-culture samples for each natural haplotype, for a total of eight pyrosequencing reactions using cDNA and eight pyrosequencing reactions using gDNA for each of the 48 strains (Supplementary Table 3).

In gDNA samples from hybrids, the two *TDH3* alleles are expected to be equally abundant; however, differences in PCR amplification of the two alleles (or aneuploidies altering copy number of *TDH3*) can cause unequal representation in the pyrosequencing data. Because such deviations cause estimates of relative allelic expression for these samples to be less reliable, the 15% of samples with gDNA ratios that deviated by more than 15% from the expected 50:50 ratio were excluded. Relative abundance of the two *TDH3* alleles is expected to be more variable in the co-cultured samples because of unequal representation from differences in concentration of the two genotypes before mixing and/or after growth. Samples from co-cultured diploids with gDNA ratios in the upper or lower 10th percentiles were also excluded from analysis. These quality control filters left 48 strains with at least two replicates in both the hybrid and co-cultured samples.

For each sample, relative allelic abundance in the cDNA sample was divided by relative allelic abundance for the corresponding gDNA sample to correct for remaining biases⁴¹. These ratios (Y_{ijk}) from strain i , plate j , and replicate k were fitted to the following linear model, including strain (ranging from 1 to 48) and plate (ranging from 1 to 3) as fixed effects as well as the cell density of the sample before and after growth from which the RNA and DNA were extracted (measured by $D_{600\text{ nm}}$) as a covariate: $Y_{ijk} = \mu + \text{strain} + \text{plate} + \text{density.0} + \text{density.1} + \varepsilon$. An analysis of variance (ANOVA) found that strain, plate, and initial density were statistically significant for hybrids (strain: $P = 1.38 \times 10^{-20}$; plate: $P = 1.01 \times 10^{-10}$; density.0: $P = 5.01 \times 10^{-3}$; density.1: $P = 0.740$), and strain and plate were statistically significant for co-cultured diploids (strain: $P = 8.16 \times 10^{-20}$; plate: $P = 2.65 \times 10^{-3}$; density.0: $P = 0.734$; density.1: $P = 0.833$). Expression values for each sample were adjusted to remove the effects of plate and initial cell density. Differences in allelic abundance caused by the synonymous change introduced for pyrosequencing were estimated by analysing a hybrid between BY4741 and YPW360 and a co-culture of BY4741 and YPW362. The effects of this change were then subtracted from the log₂-transformed expression ratio for all samples. Strains with significant *cis*-regulatory divergence from the reference were identified using t -tests. R code used for these analyses is provided in Supplementary File 2.

To determine the amount of variation in *TDH3* *cis*-regulatory activity explained by strain identity and the *TDH3* promoter haplotype, we fitted the normalized expression values to linear models containing fixed effects of either strain identity or promoter haplotype alone. Variance among strains explained by strain identity was assumed to reflect heritable variation, with residual variance assumed to result from technical noise. Because multiple strains contained the same *TDH3* promoter haplotype, we were able to determine the proportion of this heritable variance explained by polymorphisms in the *TDH3* promoter region tested. Seventy-five per cent of all *cis*-regulatory variation and 97% of heritable *cis*-regulatory variation were explained by the *TDH3* promoter haplotype. To estimate the error associated with these estimates of variance explained, we analysed 100,000 bootstrap replicates of the data with the same linear models.

Constructing strains with mutations and polymorphisms in *P_{TDH3}*. To assay *cis*-regulatory activity of the *TDH3* promoter efficiently, we used a *P_{TDH3}*-YFP reporter gene integrated near a pseudogene on chromosome 1 of strain BY4724 at position 199270 (ref. 9). This *P_{TDH3}*-YFP transgene contained a 678 bp sequence including the *TDH3* promoter that was fused to the coding sequence for YFP and the *CYC1* (cytochrome *c* isoform 1) terminator. The 678 bp sequence extended 5' from the start codon of *TDH3* into the 3' UTR of the neighbouring gene (*PDX1*),

including the 5' UTR of *TDH3*. To facilitate replacing this reference haplotype with other *P_{TDH3}* haplotypes, we used homologous replacement to create a derivative of this starting strain in which the *P_{TDH3}* sequence as well as the start codon of YFP was replaced with the *URA3* gene (*URA3-YFP*; strain YPW44).

To assess *cis*-regulatory activity of naturally occurring *P_{TDH3}* haplotypes, we amplified the *TDH3* promoters from the 85 natural isolates using PCR and transformed these PCR products into the *URA3-YFP* intermediate. Unobserved intermediate haplotypes between all pairs of haplotypes that differed at exactly two sites were constructed by PCR-mediated site-directed mutagenesis of one of the two haplotypes in each pair and transformed into the *URA3-YFP* strain. The 236 mutant *P_{TDH3}* alleles analysed, each containing a single G:C→A:T transition, were also constructed using PCR-mediated site-directed mutagenesis, but starting with the reference *P_{TDH3}* haplotype. Each of these sequences was also transformed into the same *URA3-YFP* strain. All PCR primers used for amplification and site-directed mutagenesis are shown in Supplementary Table 2. In all cases, (1) transformations were performed using the lithium acetate method⁴²; (2) transformants were selected on 5-FOA plates, streaked for single colonies, and confirmed to not be petite (missing mitochondrial DNA) by replica plating onto YPG plates (3% (v/v) glycerol, 2% Oxoid yeast extract, 2% Oxoid peptone, 2% agar); and (3) Sanger sequencing was used to determine the sequence of potential transformants.

Quantifying fluorescence of *P_{TDH3}-YFP*, a proxy for *cis*-regulatory activity of *P_{TDH3}*. Prior work shows that fluorescence of reporter proteins such as YFP provide a reliable readout of *cis*-regulatory activity^{9,43}. Before quantifying fluorescence, all strains were revived from glycerol stocks onto YPG at the same time to control for age related effects on expression. Strains were inoculated from YPG solid media into 500 µl of YPD liquid media and grown for 20 h at 30 °C in 2 ml 96-well plates with 3 mm glass beads, shaking at 250 rpm. Immediately before flow cytometry, 20 µl of the overnight culture was transferred into 500 µl of SC-R (dextrose) media⁹. Flow cytometry data were collected on an Accuri C6 using an Intellicyt Hypercyt Auto-sampler. Flow rate was 14 µl min⁻¹ and core size was 10 µm. A blue laser (λ = 488 nm) was used for excitation of YFP. Data were collected from FL1 using a 533/30 nm filter. Each culture was sampled for 2–3 s, resulting in approximately 20,000 recorded events.

Samples were processed using the flowClust⁴⁴ and flowCore⁴⁵ packages within R (version 3.0.2) and custom R scripts⁴⁶ (<http://www.r-project.org/>) (Supplementary File 3). Raw data (Extended Data Fig. 8a) were log₁₀ transformed and artefacts were removed by excluding events with extreme FSC.H, FSC.A, SSC.H, SSC.A, and width values (Extended Data Fig. 8b). Samples were clustered on the basis of FSC.A and width to remove non-viable cells and cellular debris, then clustered on FSC.H and FSC.A to remove doublets (Extended Data Fig. 8c). Finally, samples were clustered on FL1.A and FSC.A to obtain homogeneous populations of cells in the same stage of the cell cycle (Extended Data Fig. 8d). At each filtering step, data were divided into exactly two clusters. Samples containing fewer than 1,000 events after processing were discarded. For each sample, YFP expression was calculated as the median log₁₀(FL1.A)²/log₁₀(FSC.A)³. This corrected YFP expression levels for the correlation between fluorescence and cell size (measured by FSC.A) (Extended Data Fig. 8e). Expression noise for each sample was calculated as σ/μ . The following alternative metrics for expression noise were also calculated and used for analysis: σ , σ^2/μ^2 , σ^2/μ , and residuals from a regression of σ on μ .

For each genotype, nine independent replicate cultures were analysed, with three biological replicates included on each of three different days. Power analyses indicated that six replicates were sufficient to detect differences in mean expression of 2%; at $\alpha = 0.05$ and power greater than 90%. To control for variation in growth conditions, all plates contained 20 replicates of the wild-type reference strain, with at least one control sample in each row and column of the plate. For both mean expression and the standard deviation of expression, the control samples were fitted to a linear model that included final cell number and average cell width as well as the day, replicate, array, read order, growth position in the incubator, array depth in incubator, measurement block, row, and column of the sample. Stepwise Akaike information criterion was performed on this model to identify the most informative combination of variables to keep in the model. Plate (which incorporated effects of day, replicate, and array) and block were significant from this model. The effects of these factors were removed from measures of YFP (Extended Data Fig. 8f–y) before the final analysis. A non-fluorescent strain containing no *TDH3* promoter was used to estimate autofluorescence and this value was subtracted from all YFP expression values (Supplementary File 4 and Supplementary Table 4).

The effect of an individual polymorphism on mean expression level and expression noise was measured as the difference in phenotype between the descendant and ancestral haplotypes that varied only for that polymorphism. The effect of an individual mutation on mean expression level and expression noise was measured as the difference in phenotype between the reference strain and the strain carrying that mutation. Statistical significance of effects for individual polymorphisms and mutations was assessed using two-sided *t*-tests.

Although we frequently switched to fresh clones from glycerol stocks of the *URA3-YFP* strain during construction of the collection of 381 *P_{TDH3}-YFP* strains analysed in this study, we checked for the presence of relevant second-site mutations that might have arisen spontaneously by independently reintroducing the *P_{TDH3}* reference allele three times. No difference in YFP fluorescence was observed among these replicate strains for either mean expression level or expression noise (mean $P = 0.16$, noise $P = 0.069$, $n = 1,483$, ANOVA).

The reference haplotype used to determine the effect of new mutations differs from the most closely related natural haplotype (haplotype A) by a single base pair. To determine the impact of this single nucleotide difference on the distribution of mutational effects for mean expression level and expression noise, we introduced 28 of the G:C→A:T mutations into haplotype A and constructed *P_{TDH3}-YFP* strains that carried these alleles. The 28 mutations chosen for testing showed a range of effects on both mean expression level and expression noise. We found that this single base difference significantly decreased mean expression level by 3.7% ($P = 8.1 \times 10^{-56}$, ANOVA) and significantly increased expression noise by 6.8% ($P = 1.61 \times 10^{-4}$, ANOVA), but these effects were largely consistent across genetic backgrounds, indicating little and/or weak epistasis (Extended Data Fig. 9a, b). Indeed, we found that the distributions of mutational effects estimated by these 28 mutations on haplotype A and the 236 mutations on the reference haplotype were similar for both mean expression level and expression noise (Extended Data Fig. 9c, d).

The reference background also contained 6 bp at the 5' end of the *P_{TDH3}* region derived from the 3' UTR of *PDX1* that was not included in the *P_{TDH3}-YFP* constructs containing natural *P_{TDH3}* haplotypes. To determine whether this sequence was likely to have affected our measurements of polymorphism effects, we tested for a significant change in YFP fluorescence when these 6 bp were added to the *P_{TDH3}-YFP* alleles carrying the natural haplotypes A, D, and VV. We found no significant difference between genotypes with and without this 6 bp sequence (mean $P = 0.88$, noise $P = 0.25$, ANOVA).

To determine the sensitivity of our conclusions to the specific genetic background used to assay *cis*-regulatory activity, we created hybrids between one of the natural *S. cerevisiae* isolates (YPS1000) and (1) 111 strains with mutations in *P_{TDH3}-YFP*, (2) the strain carrying the reference *P_{TDH3}-YFP* allele, (3) 39 strains with naturally occurring *TDH3* promoter haplotypes driving YFP expression, and (4) a strain without the *TDH3* promoter in the *P_{TDH3}-YFP* construct and thus no YFP expression. YPS1000 was isolated from an oak tree and is substantially diverged from strain BY4724 (>53,000 SNPs, 0.44% (refs 10, 11)). We crossed all 152 of the strains described above (mating type α) to an isolate of YPS1000 that contained a KanMX4 drug resistance marker at the HO locus (mating type α). Hybrids were created by mixing equal cell numbers in liquid YPD and growing at 30 °C for 48 h without shaking. Cultures were diluted and plated on YPG + G418 to select for hybrids and prevent petite cells from growing. Colonies were grown for 48 h and then screened by fluorescence microscopy for YFP expression. Fluorescent colonies were streaked for single colonies and then a single colony was randomly chosen from each plate, transferred to a new plate, and confirmed to be diploid using a PCR reaction that genotyped the mating type locus. Four replicates of each strain were arrayed as in the original experiment with 20 controls per 96-well plate. Samples were grown for 20 h in 500 µl of YPD liquid with shaking at 30 °C, then analysed using the same flow cytometer machine and conditions described above. Samples were processed using the same analysis scripts described above, and mean expression level and expression noise were calculated. Eight of the 111 genotypes carrying reporter genes with mutations as well as four of the 39 genotypes carrying reporter genes with polymorphisms showed phenotypes suggesting that they were aneuploidies. This rate was consistent with our previous observations of spontaneous aneuploidies produced by BY4742 (ref. 9). One additional strain (containing a mutation in the *TDH3* promoter) was also excluded for having highly inconsistent measurements among replicate populations. The R script used for this analysis is provided in Supplementary File 5 and the data are provided in Supplementary Table 5.

Tests for evidence of natural selection. In the absence of selection, the effects of polymorphisms are expected to be consistent with the effects of a random sample of new mutations. Because our data were non-normally distributed, we used non-parametric tests based on sampling to assess significance. To estimate the probability of occurrence for a mutation with a particular effect (x), we used a Gaussian kernel with a bandwidth of 0.01 to fit density curves to the distributions of mutational effects observed both for mean expression level and for expression noise. We calculated the density for mean expression level values ranging from 0% to 200%, and for expression noise values ranging from 0% to 800%, ranges that extended beyond all observed effects. We set the minimum density for any effect size to $1/(\text{number of mutations included in the mutational distribution})$. We expected this minimum to overestimate the true probability of most unobserved effect sizes, making this a conservative baseline for testing whether the effects of observed polymorphisms were a biased subset of all possible mutations. These density curves

were then converted into probability distributions by setting the total density equal to 1 (Extended Data Fig. 10a, b).

To calculate the log-likelihood of a set of n genetic variants with effects x_1, x_2, \dots, x_n , we used these probability distributions to estimate the log-likelihood of a mutation with that effect, $p(x)$, and summed probabilities for all genetic variants. That is, the log-likelihood of a set of particular effects was calculated as $\sum_{i=1}^n \log(p(x_i))$. The log-likelihood calculated for the 45 observed polymorphisms was compared with the log-likelihoods of 100,000 samples of 45 mutations drawn randomly from the corresponding mutational distribution with replacement. To test the hypothesis that the effects of observed polymorphisms were unlikely to result by chance from the mutational process alone, one-sided P values were calculated as the proportion of random samples with log-likelihoods less than the log-likelihood value calculated for the observed polymorphisms. To determine the effects of mutations in the known TFBS on this test for selection, we excluded the effects of the mutations in the known TFBS from the distribution of mutational effects, recalculated the density curves and probability distributions, then recalculated the log-likelihoods and P values.

Fitness functions relate the effect of a new mutation to its likelihood of survival within a population. We determined the most likely fitness function for mean expression level and expression noise by using a hill-climbing algorithm to identify the α and β parameters of a beta distribution that maximized the likelihood of the observed polymorphism data when multiplied by the distribution of mutational effects. The beta function was started with parameters consistent with neutral evolution ($\alpha = 0, \beta = 0$) and new parameters were sampled randomly from a uniform distribution. The likelihood of the observed data was then calculated under the combined distribution of mutational effects and the new beta distribution. If the likelihood increased, the new parameters were kept; if not, they were discarded. This process was repeated until we observed 1,000 successive rejections. After each rejection, the width of the uniform distribution was increased to sample values farther away from the current parameters. A likelihood ratio test (two degrees of freedom) comparing the fitness function described by the maximum likelihood parameters for the beta distribution with a fitness function consistent with neutrality ($\alpha = 0, \beta = 0$) was used to test for statistically significant evidence of selection.

If the effects of polymorphisms are determined solely by mutation, phenotypes should drift over evolutionary time in a manner dictated by the mutational process. We modelled such a neutral scenario by starting with the phenotype of the inferred common ancestor and adding to it effects randomly drawn from the mutational distribution (sampled with replacement) for each new polymorphism observed in the haplotype network, maintaining the observed relationships among haplotypes. This process was repeated 10,000 times to generate a range of potential outcomes consistent with neutral evolution of P_{TDH3} activity. We then compared the observed polymorphism data with the results of these neutral simulations to test for a statistically significant deviation from neutrality that would indicate selection. A more detailed description of this method follows.

Let x be the number of new polymorphisms added to the population to convert an observed haplotype into the most closely related descendent haplotype in each lineage that exists or must have existed in wild populations of *S. cerevisiae*. In the haplotype network for P_{TDH3} , x ranges from 0 to 5 (Extended Data Fig. 2a). Pairs of haplotypes separated by 0 new polymorphisms result from recombination between existing haplotypes (for example, haplotype RR, which is a recombinant of haplotypes W and FF).

The probability of a polymorphism with any particular effect being added to the population was assumed, in the absence of selection, to be equal to the probability of a new mutation with that effect. The log-likelihood of a single mutation ($x = 1$) with a particular effect was calculated using the probability distributions fitted to density curves based on the observed mutational distributions described above. To generate equivalent probability distributions for sets of $x = 2, 3, 4$, or 5 new mutations, we randomly drew x mutations from the observed distribution of single mutational effects with replacement, calculated the combined effect of these mutations, and repeated this process 10,000 times. We then fitted a density curve to these 10,000 combined effect values for each value of x , set the total density to 1 to convert this into a probability distribution, and used these curves (Extended Data Fig. 10c, d) to calculate the log-likelihood of a particular set of x new polymorphisms with a given combined effect in the absence of selection. A likelihood of 1 was assigned to pairs of haplotypes separated only by recombination ($x = 0$), because the new genetic variant incorporated into the descendant haplotype was already known to have arisen in the population.

To calculate an overall log-likelihood for the observed set of polymorphisms, we summed the log-likelihood values for phenotypic differences observed between each pair of most closely related haplotypes seen among the natural isolates. To determine whether this overall log-likelihood for the observed polymorphisms was consistent with neutrality, we used the structure of the haplotype network to simulate 10,000

alternative sets of haplotype effects assuming that the effect of each new polymorphism was drawn randomly from the distribution of mutational effects. We calculated the log-likelihood for each node, in each set of haplotype effects, as $\log[\prod_{i=1}^5 (n_x! \times \prod_{j=1}^{n_x} p(x_j))]$, where x is the number of mutational steps, n_x is the number of immediately descendent haplotypes that are x mutational steps away from the focal node that exist or must have existed in *S. cerevisiae* (Extended Data Fig. 2a), and $p(x_i)$ is the likelihood of the i th mutation drawn from the probability distribution based on sets of x mutations. The $n_x!$ factor accounts for all possible ways that x mutations (or polymorphisms) added to the population at any given step could have been arranged among the set of descendent haplotypes observed.

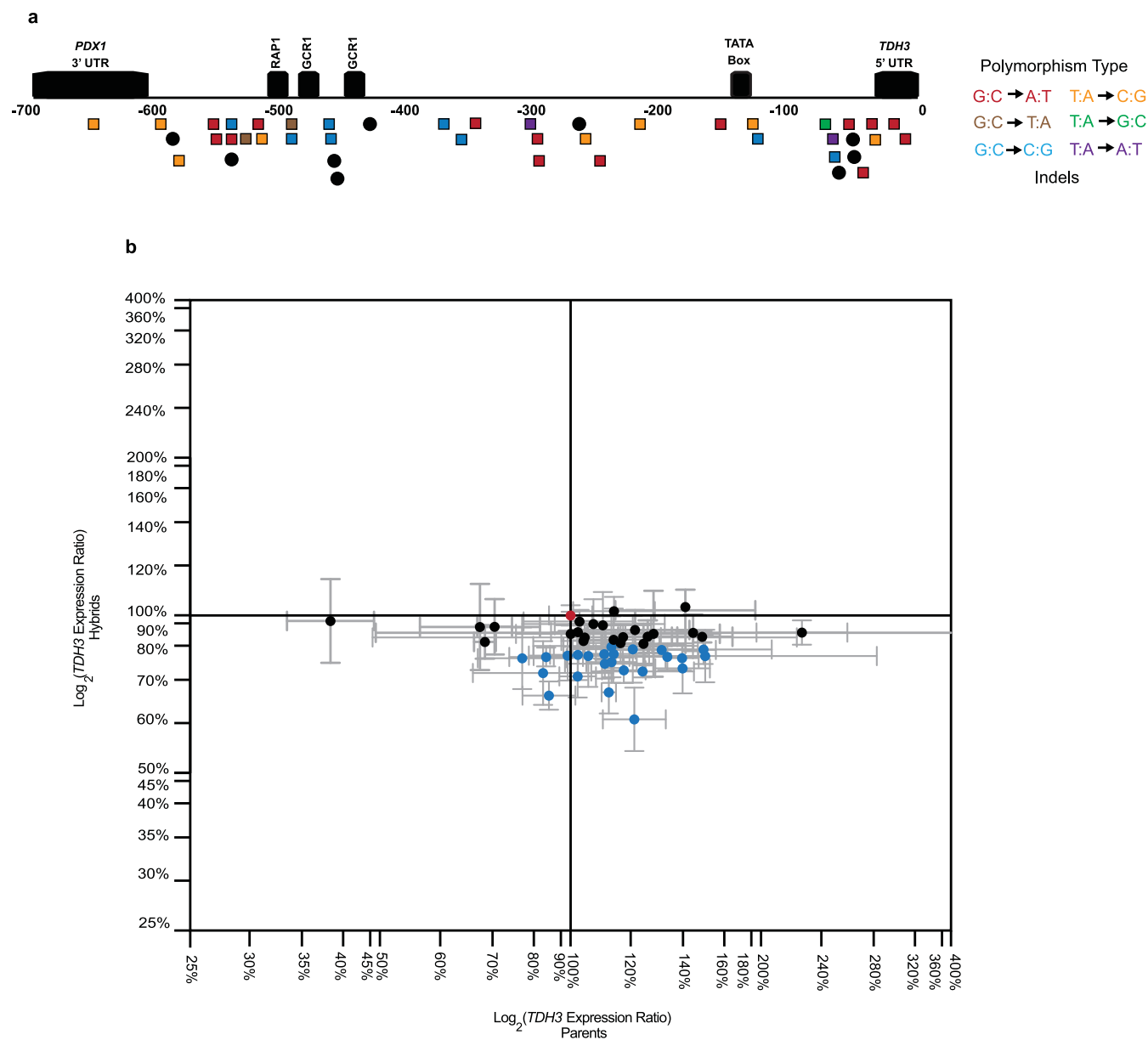
To illustrate how this works for one particularly complex node in the network, consider haplotype H and its six immediately descendent haplotypes, L, I, VV, D, S, and N (Extended Data Fig. 2a). Five of these descendent haplotypes (all except L) are all one mutational step away from H. To simulate the neutral evolution of these five haplotypes, we drew five mutational effects randomly from the probability distribution for single mutations ($x = 1$) with replacement, then determined the likelihood of each of these mutational effects based on the probability distribution for $x = 1$. These likelihood values were multiplied together to calculate the combined probability of that particular set of five mutational effects occurring. This product was then multiplied by the 5! ways in which these mutations could have been arranged among the five descendent haplotypes. We also took into account that haplotype H has one additional descendent haplotype that is five mutational steps away from H (with none of the intermediate haplotypes known) by drawing a single value randomly from the distribution of mutational effects derived from random sets of five mutations ($x = 5$); we calculated its likelihood using the probability distribution for $x = 5$; and we multiplied it by the 1! way in which this set of five mutational effects could have been added to haplotype H to produce haplotype L.

The log-likelihoods for all nodes in the haplotype network were then summed to compute the log-likelihood of each set of haplotypes. To determine whether the *cis*-regulatory phenotypes observed among the natural isolates were consistent with neutral evolution, we compared the log-likelihood calculated for the observed polymorphisms with the log-likelihoods calculated for the 10,000 data sets simulated assuming neutrality. A one-sided P value was calculated as the proportion of simulated neutral data sets that had a log-likelihood value less than the log-likelihood for the observed polymorphisms (Extended Data Figs 5g, h and 6q).

Analysis of additional mutational data sets. To test for differences in effects among different types of point mutation, we analysed data from previously published mutagenesis experiments in which the effects of individual mutations on *cis*-regulatory activity were determined^{13–16}. Effects were split into each of the 12 mutation types and plotted on the same scale for all regulatory elements (Extended Data Fig. 3). For each *cis*-regulatory element, we used an ANOVA to test for a significant difference among mutation types. In all cases, no significant effect was observed ($P > 0.05$). We also used a linear model including the identity of the *cis*-regulatory element and mutation type as main effects to test for a significant difference among mutational classes for sets of *cis*-regulatory elements across studies. Again, we found no significant difference among different types of mutation ($P = 0.68$, ANOVA).

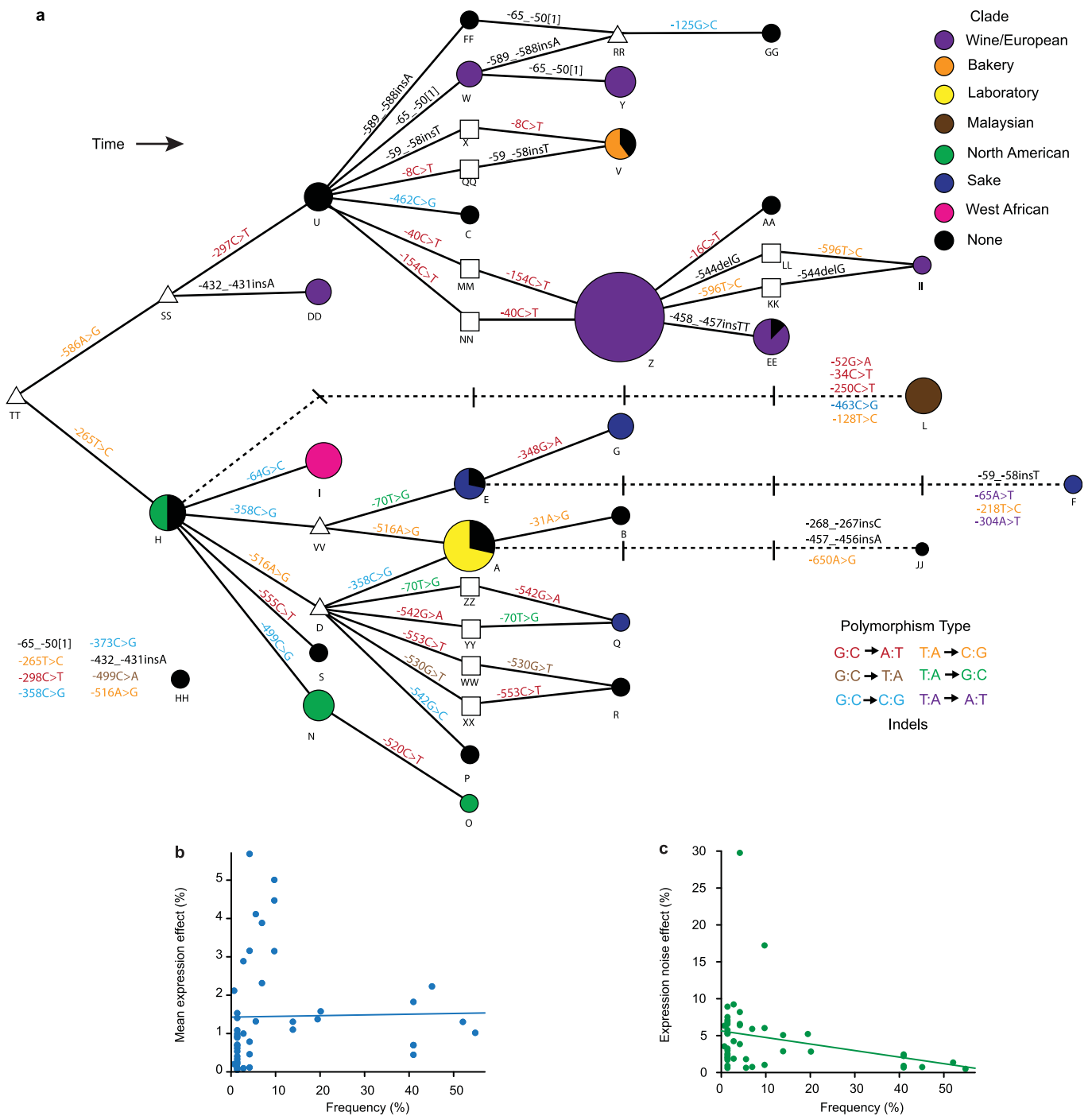
30. Taly, J.-F. *et al.* Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures. *Nature Protocols* **6**, 1669–1682 (2011).
31. Löytynoja, A. & Goldman, N. webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinform.* **11**, 579 (2010).
32. Libkind, D. *et al.* Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc. Natl Acad. Sci. USA* **108**, 14539–14544 (2011).
33. Scannell, D. R. *et al.* The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3* **1**, 11–25 (2011).
34. Liti, G. *et al.* High quality *de novo* sequencing and assembly of the *Saccharomyces uvarum* genome. *BMC Genom.* **14**, 69 (2013).
35. Wang, Q.-M., Liu, W.-Q., Liti, G., Wang, S.-A. & Bai, F.-Y. Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. *Mol. Ecol.* **21**, 5404–5417 (2012).
36. Tamura, K., Stecher, G., Peterson, D., Filipi, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
37. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659 (2000).
38. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
39. Hittinger, C. T. *Saccharomyces* diversity and evolution: a budding model genus. *Trends Genet.* **29**, 309–317 (2013).
40. Wittkopp, P. J., Haerum, B. K. & Clark, A. G. Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430**, 85–88 (2004).
41. Wittkopp, P. J. in *Molecular Methods for Evolutionary Genetics* Vol. 772 (eds Orgogozo, V. & Rockman, M. V.) 297–317 (Humana, 2011).
42. Gietz, R. & Woods, R. in *Methods in Molecular Biology* 2nd edn, Vol. 313 (ed. Xiao, W.) 107–120 (Springer, 2006).

43. Kudla, G., Murray, A., Tollervey, D. & Plotkin, J. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
44. Lo, K., Hahne, F., Brinkman, R. R. & Gottardo, R. flowClust: a Bioconductor package for automated gating of flow cytometry data. *BMC Bioinform.* **10**, 145 (2009).
45. Hahne, F. *et al.* flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinform.* **10**, 106 (2009).
46. R Core Team. R: A language and environment for statistical computing (2013).



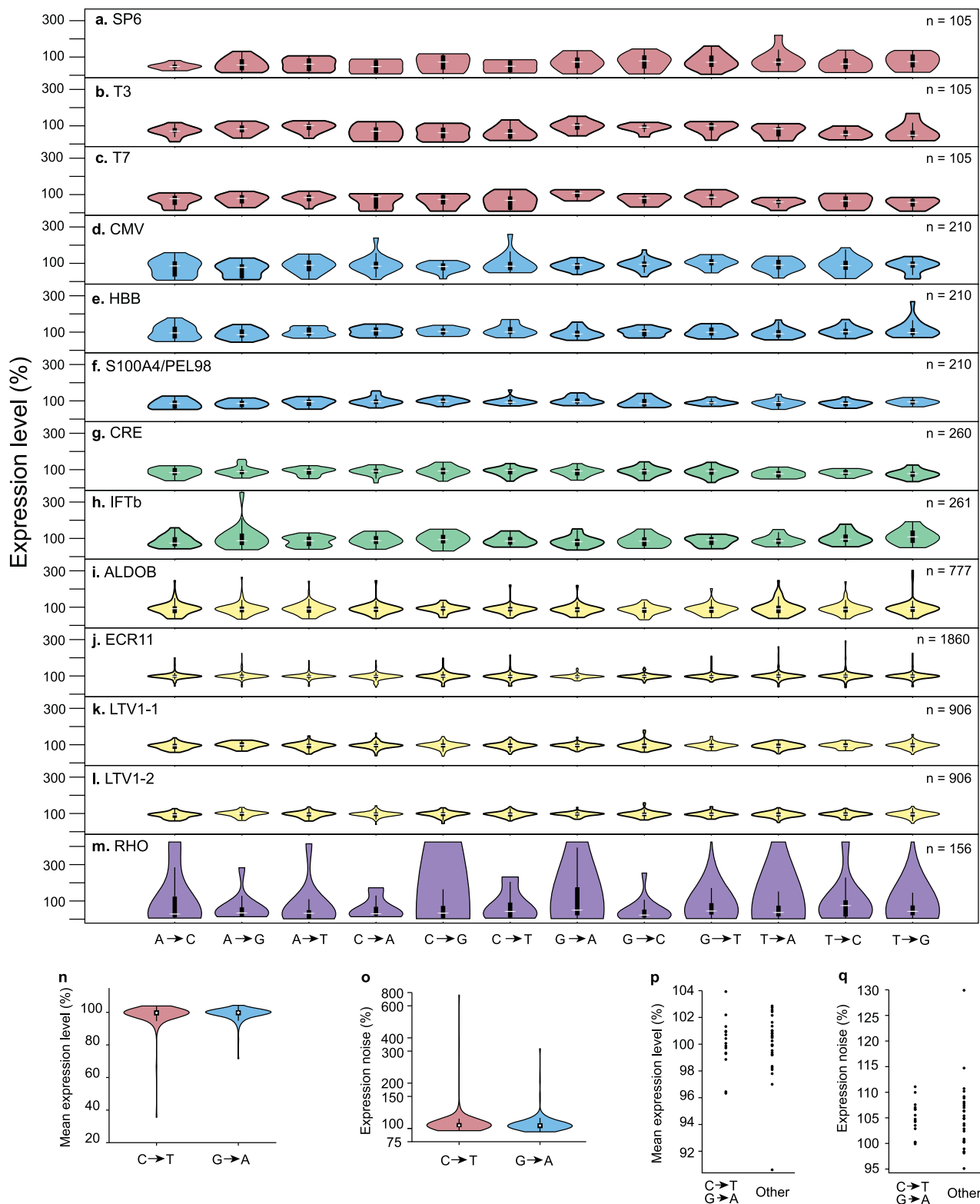
Extended Data Figure 1 | *TDH3* promoter polymorphisms influence *TDH3* mRNA levels. **a**, Locations of polymorphisms within the *TDH3* promoter relative to known functional elements, including RAP1 and GCR1 transcription factor binding sites, are shown. Squares, point mutations; circles, indels. Red, G:C→A:T; yellow, G:C→T:A; blue, G:C→C:G; orange T:A→C:G; green, T:A→G:C; purple, T:A→A:T. **b**, The \log_2 ratio of total expression divergence between natural isolates and a reference strain (x axis) versus the

\log_2 ratio of total *cis*-regulatory expression divergence between natural isolates and the reference strain (y axis). Error bars, 95% confidence intervals. The 25 of 48 strains with significant *cis*-regulatory differences from the reference strain are shown in blue. Reference strain is shown in red. These data show differences in *cis*- and *trans*-regulation among strains, but do not reveal the evolutionary changes that give rise to these differences.



Extended Data Figure 2 | Ancestral state reconstruction of the TDH3 promoter. **a**, The *TDH3* promoter haplotype network is shown with the inferred ancestral strain at the left. Circles represent haplotypes observed among the 85 strains, with their diameters proportional to haplotype frequency. The haplotypes are coloured according to clade (Supplementary Table 1). Triangles are haplotypes that were not observed among the strains sampled, but must exist or have existed as intermediates between observed haplotypes. Squares are possible intermediates connecting two observed haplotypes, but it

is unknown which of these actually exists or existed in *S. cerevisiae*. Solid lines connect haplotypes that differ by a single mutation; dashed lines connect haplotypes that differ by multiple mutations. Mutations on each branch are coloured by the mutation type as in Extended Data Fig. 1a. **b**, Relationship between the effect of a polymorphism on mean expression level and the frequency of that polymorphism among the strains sampled ($P = 0.43$). **c**, Relationship between the effect of a polymorphism on expression noise and the frequency of that polymorphism among the strains sampled ($P = 0.0028$).

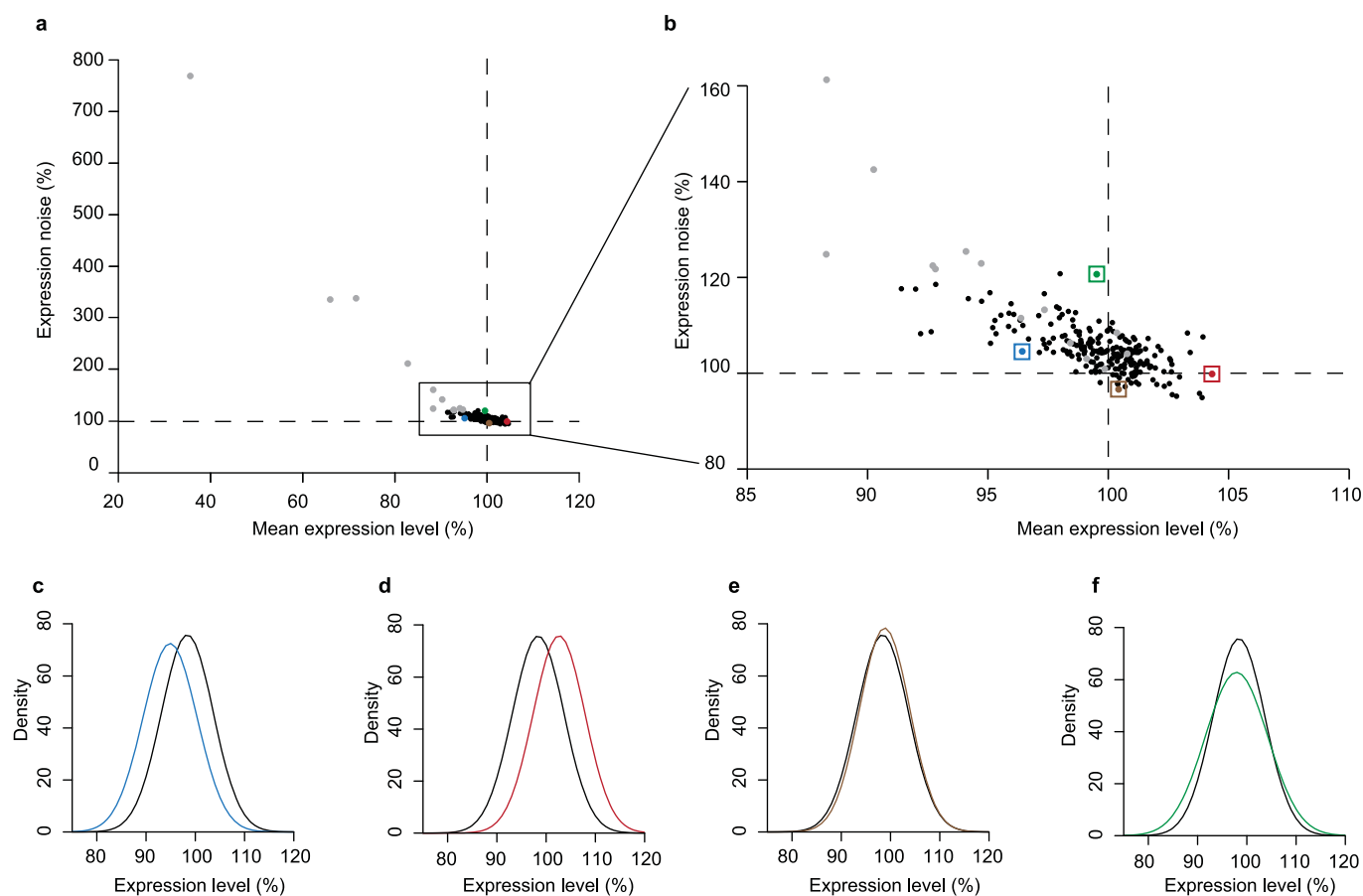


Extended Data Figure 3 | No significant difference between mutation types.

Distributions of effects on mean expression level from previous random mutagenesis experiments are shown partitioned by mutation type. For each mutation type, the distribution (inside) and density (outside, coloured) of the effects on mean expression level are shown. The number of mutations tested for each promoter is shown in the upper right corner of each panel. **a**, Bacteriophage SP6 promoter. **b**, Bacteriophage T3 promoter. **c**, Bacteriophage T7 promoter. **d**, Human CMV promoter. **e**, Human HBB promoter. **f**, Human S100A4/PEL98 promoter. **g**, Synthetic cAMP-regulated enhancer. **h**, Interferon- β enhancer. **i**, ALDOB enhancer. **j**, ECR11 enhancer.

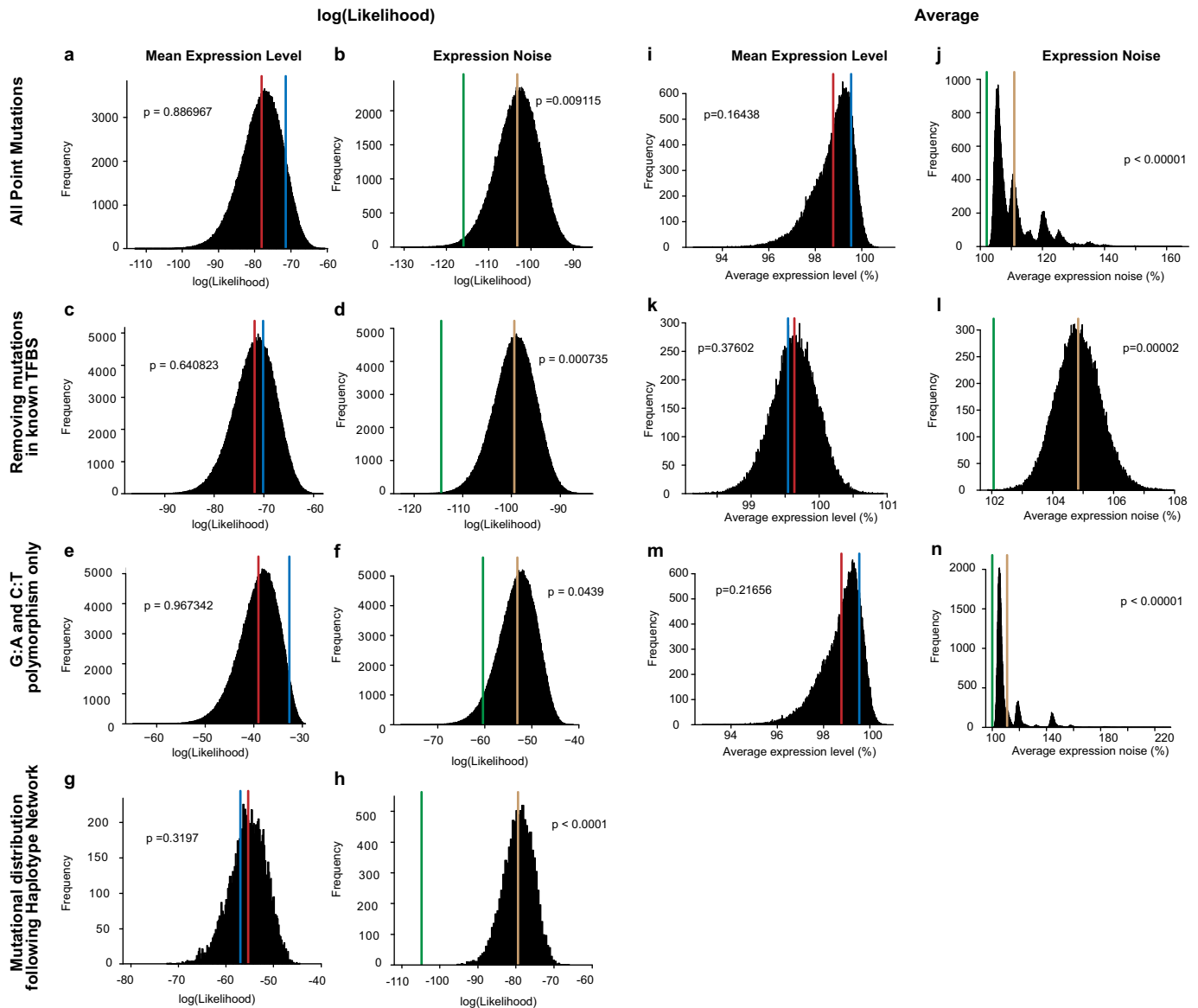
k, LTV1 enhancer replicate 1. **l**, LTV1 enhancer replicate 2. **m**, Rhodopsin promoter. Red: bacteriophage promoters from ref. 13. Blue: mammalian promoters from ref. 13. Green: mammalian enhancers from ref. 14. Yellow: mammalian promoters from ref. 15. Purple: promoter from ref. 16.

n, Distribution of effects for C → T (red) and G → A (blue) mutations for mean expression level in this study. **o**, Same as **n**, but for expression noise. **p**, Distribution of effects for C → T/G → A polymorphisms compared with other polymorphism types for mean expression level in this study. **q**, Same as **p**, but for gene expression noise.



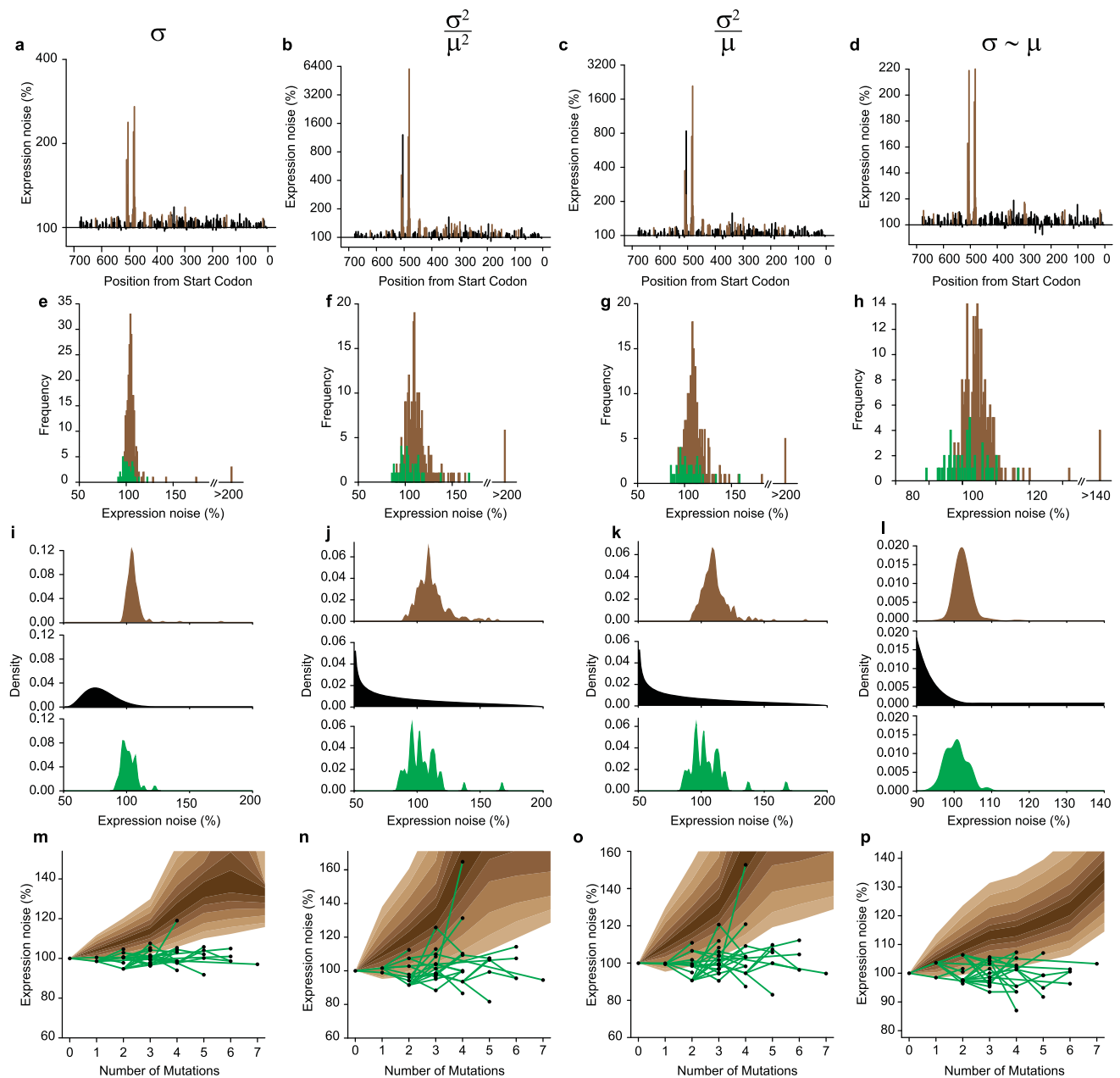
Extended Data Figure 4 | Correlation between mean expression level and expression noise. **a**, Correlation between mean expression level (x axis) and expression noise (y axis) for the 236 point mutations in the *TDH3* promoter ($R^2 = 0.85$). Grey points correspond to mutations in known transcription factor binding sites. Coloured points correspond to individual mutations highlighted in **c–f**. **b**, Alternative plot showing the majority of data from **a** more clearly; grey and coloured points are the same as in **a**. **c**, Distribution of gene expression phenotypes from a mutant (blue) with decreased mean expression level but similar expression noise as the reference strain (black). Outside the known TFBS, 50% of mutations decreased mean expression. **d**, Distribution of

gene expression phenotypes from a mutant (red) with increased mean expression level but similar gene expression noise as the reference strain (black). Outside the known TFBS, 50% of mutations increased mean expression. **e**, Distribution of gene expression phenotypes from a mutant (brown) with decreased gene expression noise but similar mean expression level as the reference strain (black). Outside the known TFBS, 13% of mutations decreased expression noise. **f**, Distribution of gene expression phenotypes from a mutant (green) with increased gene expression noise but similar mean expression level as the reference strain (black). Outside the known TFBS, 87% of mutations increased expression noise.



Extended Data Figure 5 | Tests for selection. **a–h**, Tests for selection using likelihood. **a**, The distribution of likelihood values for 100,000 randomly sampled sets of 45 mutations drawn from the mutational effect distribution is shown for mean expression level. The average likelihood for all samples of mutations tested (red) as well as the likelihood of the observed polymorphisms (blue) are also shown. **b**, Same as **a**, but for expression noise. The average likelihood for all mutation samples tested is shown in brown and the likelihood of the observed polymorphisms is shown in green. **c**, Same as **a**, but with the large effect mutations in the TFBS removed from the mutational effect distribution used for sampling. **d**, Same as **b**, but after removing the mutations in the TFBS from the mutational effect distribution. **e**, Same as **a**, but using only G→A and C→T polymorphisms. **f**, same as **b**, but using only G→A and C→T polymorphisms. **g**, Distribution of likelihoods for 10,000 random walks along the *TDH3* promoter haplotype network using the effects from the mutational distribution. **h**, Same as **e**, but for expression noise. **i–n**, Tests for selection using average effects. **i**, The distribution of average effects for 100,000 randomly sampled sets of 45 mutations drawn from the mutational effect

distribution is shown for mean expression level (black). Polymorphisms do not have a significantly different average mean expression (blue, 99.5%) than sets of mutations (red, 98.8%; $P = 0.16438$). This figure is comparable to Extended Data Fig. 5a, but uses average effects instead of the likelihoods to test for differences in distribution between random mutations and polymorphisms. **j**, Same as **i**, but for expression noise. Polymorphisms have significantly lower average expression noise (green, 102.1%) than sets of random mutations (brown, 110.9%; $P < 0.00001$). **k**, Same as **i**, but with the large effect mutations in the TFBS removed from the mutational effect distribution used for sampling (polymorphisms, 99.5%; mutations, 99.6%; $P = 0.37602$). **l**, Same as **j**, but after removing the mutations in the TFBS from the mutational effect distribution (polymorphisms, 102.1%; mutations, 104.8%; $P = 0.00002$). **m**, Same as **i**, but using only G→A and C→T polymorphisms (polymorphisms, 99.7%; mutations, 98.8%; $P = 0.21656$). **n**, Same as **j**, but using only G→A and C→T polymorphisms (polymorphisms, 100.0%; mutations, 110.9%; $P < 0.00001$).

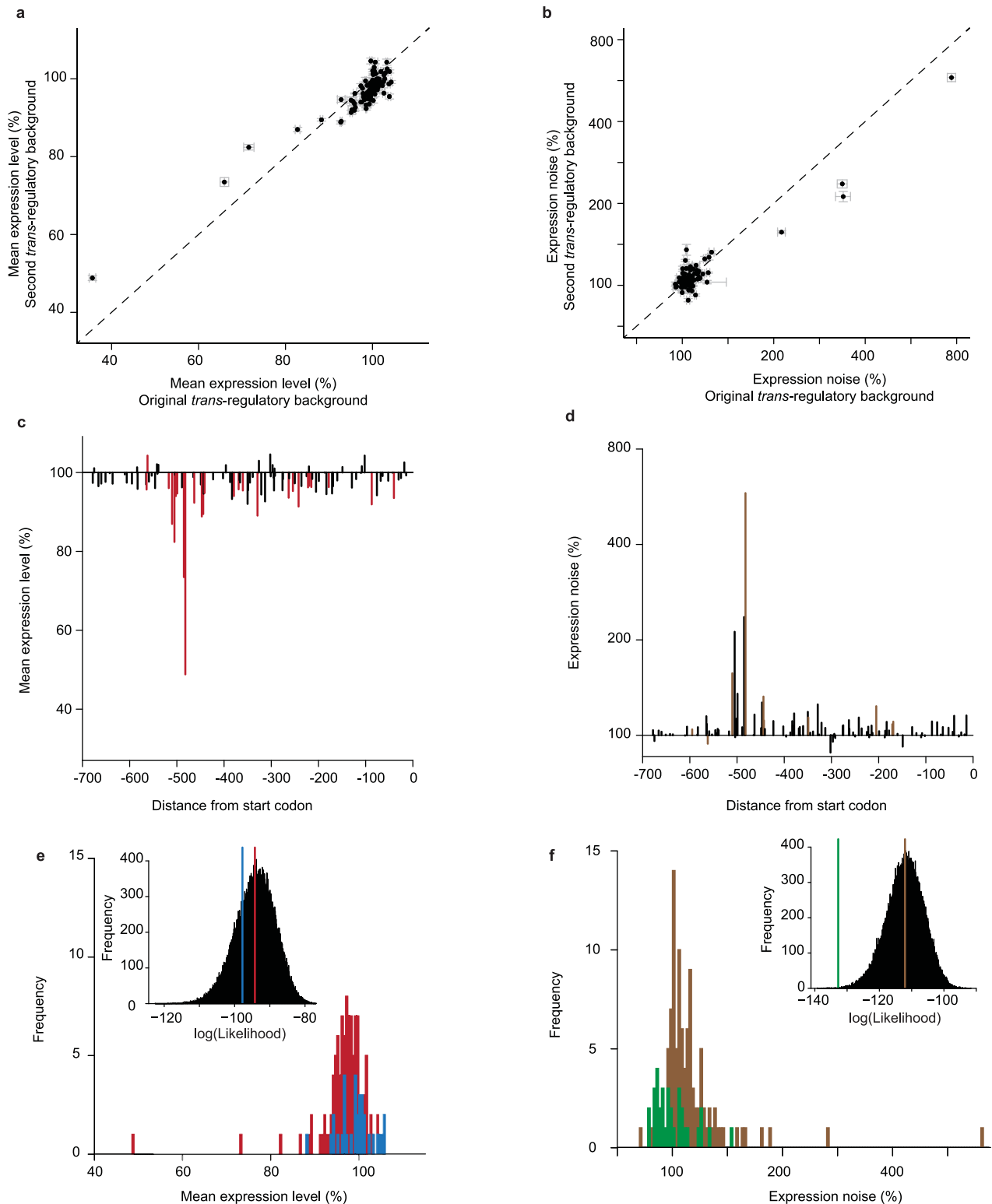


Metric [†]	Distribution [§]	No TFBS Mutations [§]	G:A/C:T Polymorphism [§]	Fitness Function [‡]	Network
μ	0.88697	0.64082	0.96734	0.87	0.3197
σ	<0.00001	<0.00001	<0.00001	9×10^{-6}	<0.0001
σ/μ	0.00912	0.00076	0.04390	0.00019	<0.0001
σ^2/μ^2	<0.00001	<0.00001	0.00760	0.00002	<0.0001
σ^2/μ	<0.00001	<0.00001	<0.00001	0.00015	<0.0001
Regression [‡]	<0.00001	<0.00001	<0.00001	4×10^{-7}	<0.0001

[†]mean expression (μ) [‡]standard deviation in expression (σ) [§]Residuals from linear model: $\sigma \sim \mu$
[§]100,000 permutations ^{||}10,000 permutations [‡]Likelihood Ratio

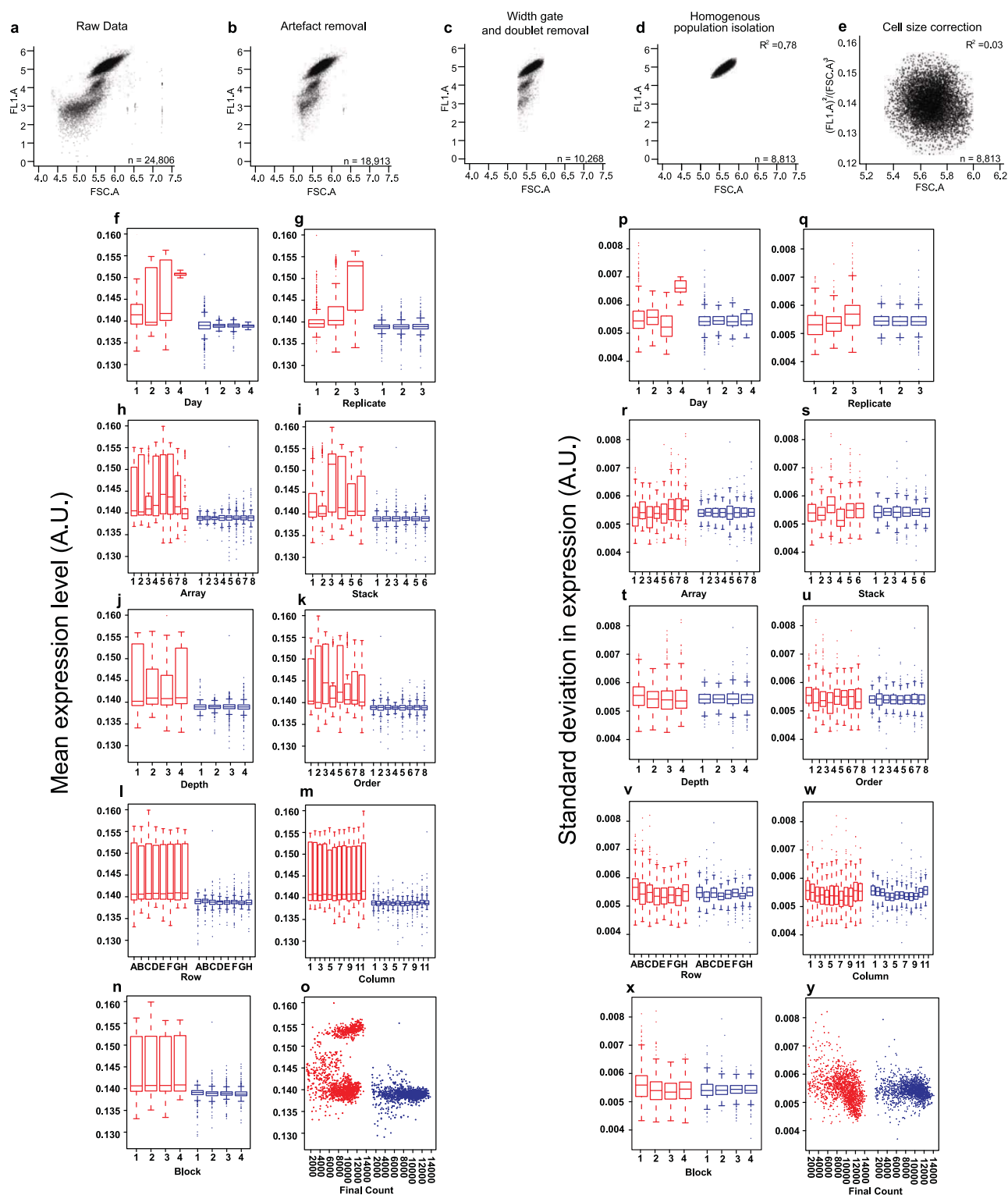
Extended Data Figure 6 | Test for selection using alternative metrics for quantifying gene expression noise. **a–d**, Distributions of effects for mutations on gene expression noise across the *TDH3* promoter with expression noise quantified as σ (**a**), σ^2/μ^2 (**b**), σ^2/μ (**c**), and residuals from the regression of σ on μ (**d**). **e–h**, Distributions of effects for mutations on gene expression noise (brown) compared with polymorphisms (green) with noise quantified as σ (**e**), σ^2/μ^2 (**f**), σ^2/μ (**g**), and residuals from the regression of σ on μ (**h**). **i–l**, The maximum likelihood fitness function (middle, black) relating the distribution of mutational effects (top, brown) to the distribution of observed polymorphisms (bottom, green) for expression noise quantified as σ (**i**),

σ^2/μ^2 (**j**), σ^2/μ (**k**), and residuals from the regression of σ on μ (**l**). **m–p**, Changes in expression noise observed among haplotypes over time in the inferred haplotype network (Extended Data Fig. 2a) are shown in green. The brown background represents the 95th, 90th, 80th, 70th, 60th, and 50th percentiles, from light to dark, for expression noise resulting from 10,000 independent simulations of phenotypic trajectories in the absence of selection where noise is quantified as σ (**m**), σ^2/μ^2 (**n**), σ^2/μ (**o**), and residuals from the regression of σ on μ (**p**). **q**, *P* values for tests of selection using mean expression (μ) and five metrics of expression noise, including σ/μ which is used throughout the main text.



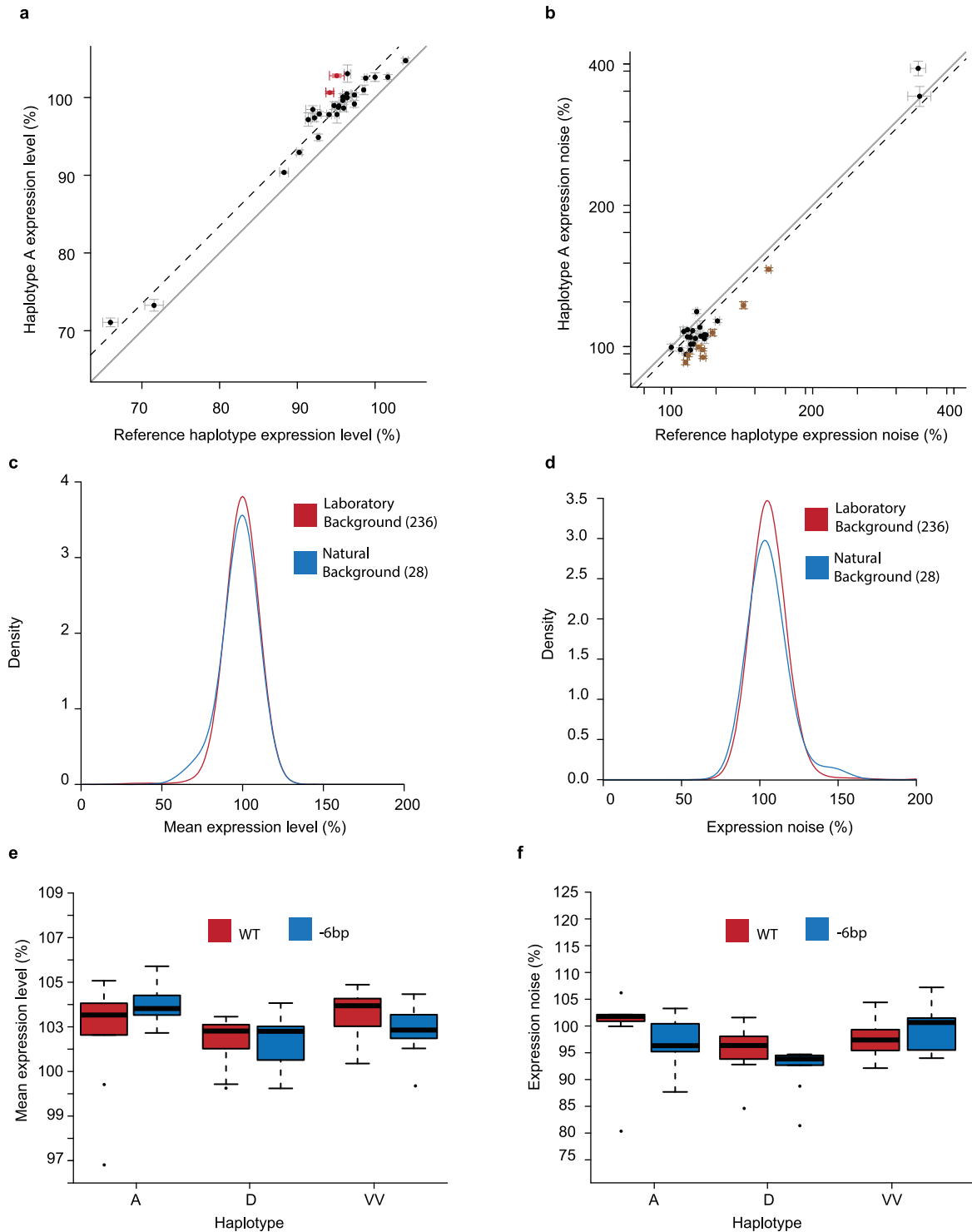
Extended Data Figure 7 | Effects of mutations and polymorphisms on a second *trans*-regulatory background. **a**, A comparison between effects of mutations on mean expression in the original *trans*-regulatory background (x axis) and a hybrid *trans*-regulatory background between BY4741 and YPS1000 (y axis). Error bars, 95% confidence intervals. **b**, Same as **a**, but for gene expression noise. **c**, Effects of individual mutations on mean expression level in the hybrid *trans*-regulatory background are shown in terms of the percentage change relative to the un-mutagenized reference allele, and are plotted according to the site mutated in the 678 bp region (significant mutations: red lines, *t*-test, Bonferroni corrected). Note that most mutations decrease expression, unlike in the original genetic background. **d**, Same as **c**, but for gene expression noise (significant mutations: brown lines, *t*-test, Bonferroni

corrected). **e**, Distribution of *de novo* mutation effects in the second *trans*-regulatory background (red) compared with the effects of naturally occurring haplotypes in this *trans*-regulatory background (blue). Inset: the distribution of likelihood values for 100,000 randomly sampled sets of 27 mutations drawn from the mutational effect distribution is shown for mean expression level. The average likelihood for all samples of mutations tested (red) as well as the likelihood of the observed polymorphisms (blue) are also shown ($P = 0.2584$). Removing mutations in the known TFBS resulted in a significant difference between mutations and polymorphisms ($P = 0.00781$). **f**, Same as **e**, but for gene expression noise. Mutations, brown. Polymorphisms, green ($P = 0.00037$). Removing mutations in the known TFBS did not change this result ($P < 0.00001$).



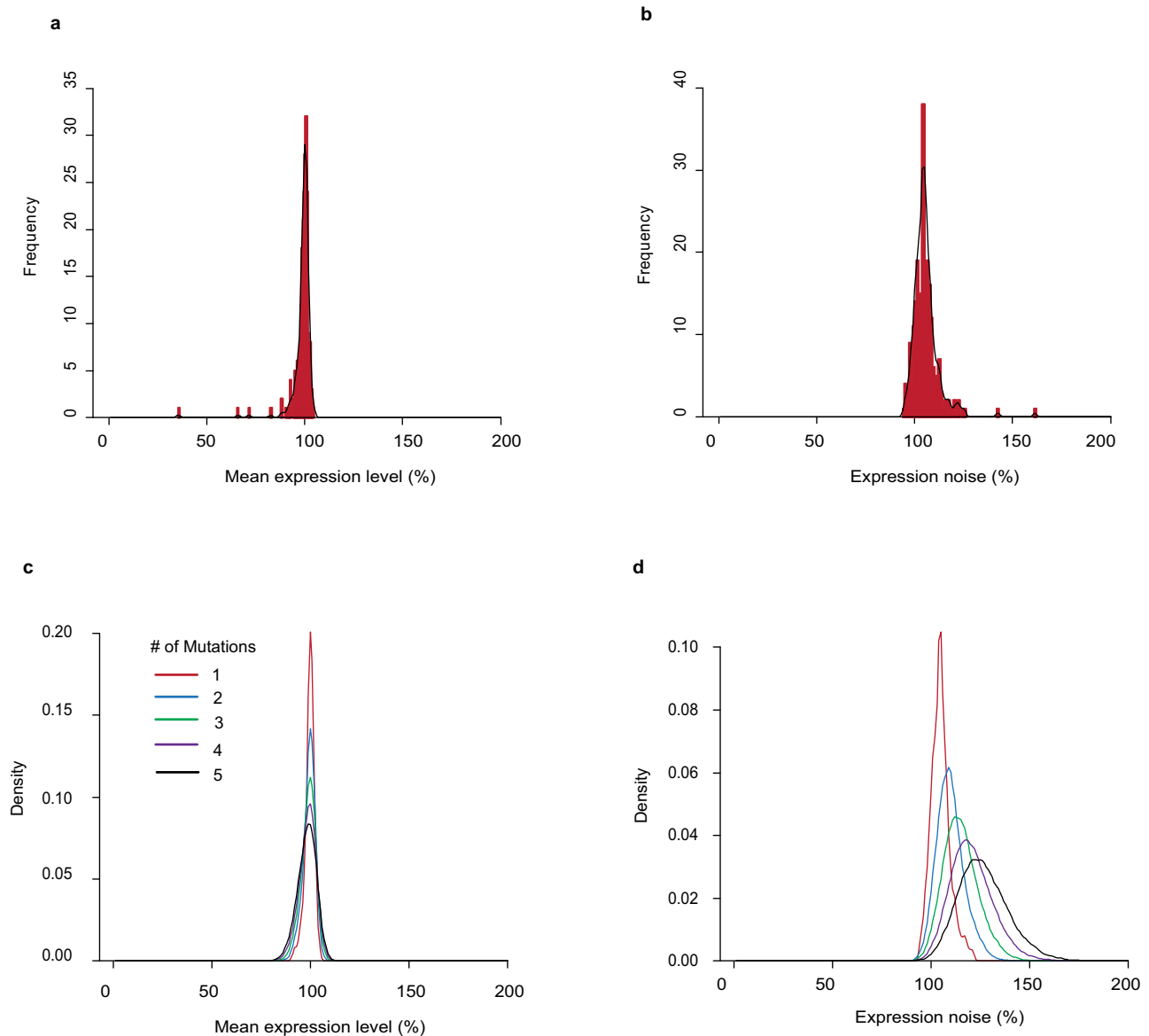
Extended Data Figure 8 | Methodology for the analysis of flow cytometry data. **a**, Raw data from the flow cytometer are shown for the first control sample collected. Each point is an individual event scored by the flow cytometer, the vast majority of which are expected to be cells. FSC.A is a proxy for cell size, and FL1.A is a measure of YFP fluorescence. Log₁₀ values are plotted both for FSC.A and for FL1.A. **b**, The same sample is shown after events found in the negative control sample (using hard gates on FSC.A and FL1.A) were excluded. **c**, The same sample is shown after flowClust was used to remove events likely to be from multiple cells entering the detector simultaneously. **d**, The same sample is shown after flowClust was used to isolate the densest homogenous population within the sample. The R^2 value shown is the correlation between YFP fluorescence and cell size. **e**, After correcting for

differences in cell size, the correlation between YFP fluorescence and cell size was nearly 0 and not significant. In all panels, the number of events analysed (that is, sample size) is shown in the bottom right corner. Box plots of mean expression of control samples before (red) and after (blue) correcting for the effects of individual plates for each day on which samples were run (**f**), for replicates nested within day (**g**), for array nested within day and replicate (**h**), for stack nested within day (**i**), for depth nested within day (**j**), for order nested within day and replicate (**k**), for row nested within array (**l**), for column nested within array (**m**), for block nested within array (**n**), and for the final cell count (**o**). The y axis is in arbitrary units. **p–x**, Same as **f–o**, but for gene expression noise.



Extended Data Figure 9 | Consistency of mutational effects on different genetic backgrounds. **a**, The effects on mean expression level for each of the 28 mutations tested on both the reference haplotype (x axis) and natural haplotype A observed in wild strains (y axis) are shown. These two haplotypes differ by a single point mutation. Solid lines show expression from the P_{TDH3} haplotypes on which the two sets of mutations were created, both of which were defined as 100% activity. The grey line shows $y = x$. The dashed line shows the consistent increase in mean expression level when these mutations were tested on haplotype A. Error bars, 95% confidence intervals. Coloured points have

significantly different effects on the two backgrounds ($P < 0.05$, ANOVA, Bonferroni corrected), indicating weak epistasis. **b**, Same as **a**, but for gene expression noise. **c**, Distributions of mutational effects for mean expression levels based on the 236 point mutations tested on the reference haplotype (red) as well as for the 28 mutations tested on haplotype A (blue). **d**, Same as **c**, but for gene expression noise. **e**, The effect on mean expression of the full *TDH3* promoter (red) compared with promoters containing six fewer base pairs at the 5' end (blue). Each box plot summarizes data from nine replicates. **f**, Same as **e**, but for expression noise.



Extended Data Figure 10 | Probability distributions for mutational effects.

a. A histogram summarizing the mutational effects on mean expression level is shown (red), overlaid with the density curve (black line) used to calculate the likelihood of an effect on mean expression level. **b.** Same as **a**, but for

expression noise. **c.** Density curves for the effects of one (red), two (blue), three (green), four (purple), or five (black) mutations randomly drawn from the distribution of mutational effects observed for mean expression level. **d.** Same as **c**, but for expression noise.

Selective corticostriatal plasticity during acquisition of an auditory discrimination task

Qiaojie Xiong^{1*}, Petr Znamenskiy^{1,2*} & Anthony M. Zador¹

Perceptual decisions are based on the activity of sensory cortical neurons, but how organisms learn to transform this activity into appropriate actions remains unknown. Projections from the auditory cortex to the auditory striatum carry information that drives decisions in an auditory frequency discrimination task¹. To assess the role of these projections in learning, we developed a channelrhodopsin-2-based assay to probe selectively for synaptic plasticity associated with corticostriatal neurons representing different frequencies. Here we report that learning this auditory discrimination preferentially potentiates corticostriatal synapses from neurons representing either high or low frequencies, depending on reward contingencies. We observe frequency-dependent corticostriatal potentiation *in vivo* over the course of training, and *in vitro* in striatal brain slices. Our findings suggest a model in which the corticostriatal synapses made by neurons tuned to different features of the sound are selectively potentiated to enable the learned transformation of sound into action.

Animals use sensory information to guide their behaviour. The neural mechanisms underlying the transformation of sensory responses into motor commands have been studied extensively using a two-alternative forced-choice task, in which subjects are trained to make a binary decision and indicate their choice by performing one of two actions. Defined brain areas have been implicated in the circuit performing this transformation in primates^{2,3} and rodents^{1,4–10}.

Striatal plasticity has been implicated in reinforcement learning^{11,12}, specifically at corticostriatal inputs^{13,14}, but the site or sites of plasticity engaged when animals learn to make appropriate decisions about sensory stimuli are not well established. We previously found that neurons in the primary auditory cortex projecting to the auditory striatum drive decisions in a two-alternative forced-choice auditory task¹ in which rats learn to associate the frequency of a complex auditory stimulus with either a left or right reward port (Fig. 1a, b). We hypothesized that plasticity of auditory corticostriatal connections encodes the association between frequency and the rewarded response.

To test this hypothesis, we developed a novel *in vivo* recording method with which we could monitor the strength of corticostriatal synapses, in a way that did not depend on the activity of cortical neurons. We used this to measure synaptic strength in single animals over multiple behavioural sessions during the course of learning. We first injected an adeno-associated virus expressing channelrhodopsin-2 (AAV-ChR2-Venus) into the left primary auditory cortex. This resulted in widespread expression of ChR2 in different cell types in the auditory cortex, including corticostriatal neurons and their axons in the striatum (Extended Data Fig. 1). We next implanted bundles of optical fibres and tetrodes into the left auditory striatum (Fig. 1c). Brief pulses of blue light delivered through the optical fibre excited the corticostriatal axons and elicited excitatory postsynaptic responses in the striatum (Fig. 1d). Because the striatum, like the CA1 region of the hippocampus, lacks recurrent excitatory connections, we reasoned that this *in vivo* ChR2-evoked local field potential response (ChR2-LFP) could serve as a measure of the strength of the corticostriatal synaptic connectivity¹⁵. The ChR2-LFP

had a stereotypic waveform consisting of an early and a late component (Extended Data Fig. 2a). Local pharmacological blockade of excitatory but not inhibitory transmission diminished the late component, indicating that it was mainly mediated by currents elicited by glutamatergic release from corticostriatal terminals (Fig. 1d and Extended Data Fig. 2c). The early component was resistant to all blockers including tetrodotoxin, suggesting that it is driven directly by light-evoked ChR2 currents in corticostriatal axons. The early component was not observed in the absence of ChR2 (Extended Data Fig. 3), indicating that it was not due to a photoelectric artefact, and its amplitude increased with increasing photostimulation (Extended Data Fig. 4). In subsequent analyses we normalized the ChR2-LFP to the amplitude of the early component (Extended Data Fig. 2a) to correct for fluctuations in the number of ChR2-expressing fibres recruited, and then used the initial slope of the second component as a measure of corticostriatal synaptic efficacy (Fig. 1d and Extended Data Fig. 2b). This metric was robust to changes in light intensity and was proportional to the intracellular excitatory postsynaptic current, indicating that it was a good measure of synaptic strength (Extended Data Fig. 4).

We used the ChR2-LFP to assess changes in the strength of corticostriatal synapses over the course of training in the cloud-of-tones task.

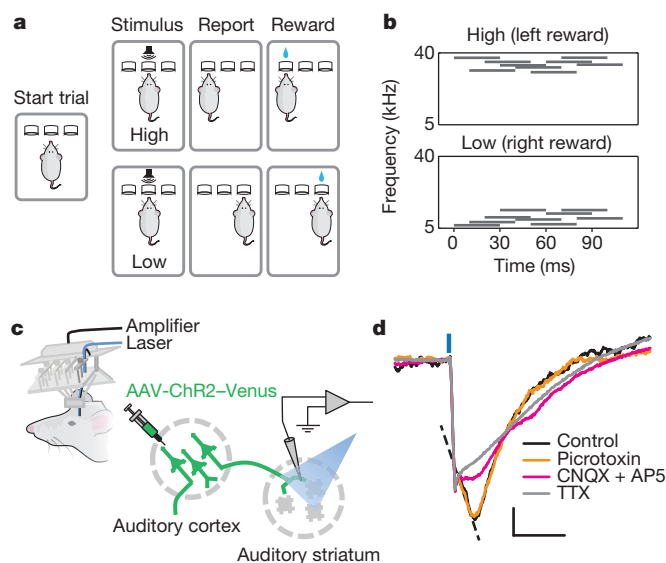


Figure 1 | Dissection of ChR2-LFP *in vivo*. **a**, Cloud-of-tones task. **b**, Example spectrograms of cloud-of-tones stimuli. **c**, Recording method to examine corticostriatal synaptic strength *in vivo*. **d**, ChR2-LFP recorded from auditory striatum under control conditions (black trace) and after application of picrotoxin (orange), CNQX (6-cyano-7-nitroquinoxaline-2,3-dione) and AP5 (2-amino-5-phosphonopentanoate) (pink), and tetrodotoxin (TTX, light grey). The slope of the CNQX/AP5-sensitive component was used to quantify corticostriatal synaptic strength (dotted line). Scale bars, 20 μ V, 5 ms.

¹Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. ²Watson School of Biological Sciences, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA.

*These authors contributed equally to this work.

After establishing a stable baseline over several days in naive rats, we measured the ChR2-LFP after each training session. We used the tone-evoked multiunit responses recorded before training to estimate the frequency tuning at each site (see Methods). At some recording sites, corticostriatal synaptic efficacy increased as soon as the animal started to learn the task, and continued to increase in subsequent training sessions (Fig. 2a). Synaptic efficacy at such sites thus reflected behavioural performance over the course of training. At other sites, however, corticostriatal efficacy remained unchanged over the course of training (Fig. 2b). We found that potentiation was restricted to sites tuned to low-frequency (<14 kHz, the centre frequency used in the task) sounds (mean potentiation 30%; $n = 17$, $P = 0.002$, signed-rank test, Fig. 2c), whereas sites tuned to high-frequency (>14 kHz) sounds showed no significant change (mean change -3%; $n = 6$, $P = 0.58$, signed-rank test, Fig. 2c.). Notably, all animals in this cohort were trained to associate low-frequency sounds with rightward choices (LowRight), and all recordings were performed in the left striatum. Hence, low frequencies were always associated with choices contralateral to the recording hemisphere. Our results therefore suggest that task training selectively enhances the strength of corticostriatal synapses only when the stimuli they encode are associated with contralateral choices (Fig. 2d).

The observed potentiation depended strongly on the frequency tuning of the recording site, suggesting that corticostriatal plasticity encodes the association of specific frequencies with rewarded actions. However, since the striatum has been widely implicated in motor learning¹⁴, we sought to rule out this and other alternative causes of plasticity unrelated to auditory discrimination. We trained animals to perform a simple two-alternative forced-choice visual task, relying on the same sequence of movements, and monitored the strength of auditory corticostriatal synapses during learning (Fig. 3a). There was no significant change in ChR2-LFP in the auditory striatum during visual task training (mean

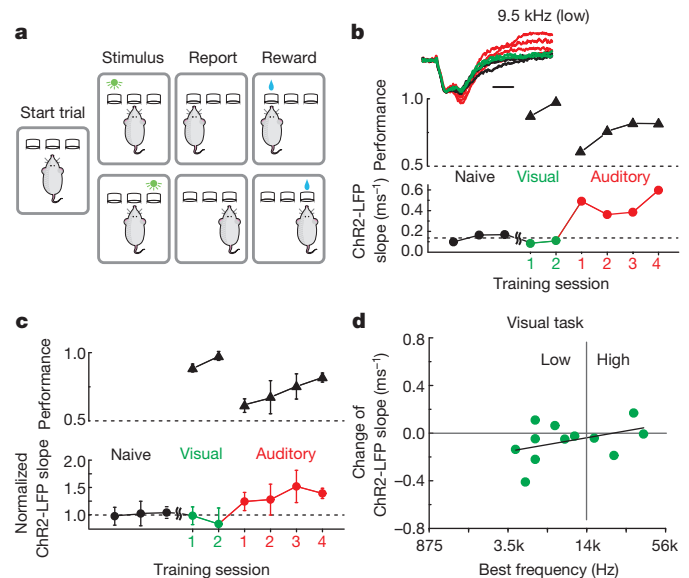


Figure 3 | Potentiation of ChR2-LFP slope is modality specific. **a**, Visual two-alternative forced-choice task. **b**, ChR2-LFP from an example auditory striatum site during visual and auditory task learning, analysed as in Fig. 2a. Scale bar, 2 ms. **c**, Population average of normalized ChR2-LFP slope during visual and auditory task training. **d**, Visual training fails to potentiate auditory striatal input (12 recording sites from four rats; least squares regression, $P = 0.192$).

change -17%; $n = 12$, $P = 0.13$, signed-rank test), and there was no correlation between potentiation and the preferred frequency at the recording site ($n = 12$, $P = 0.19$; Fig. 3d). However, corticostriatal inputs at these same recording sites were potentiated when the animals subsequently learned the auditory cloud-of-tones task (mean potentiation 36%; $n = 6$, $P = 0.03$, signed-rank test; Fig. 3b, c). We therefore conclude that the selective potentiation of auditory corticostriatal synaptic strength is specific to the acquisition of the auditory task.

The preferential potentiation *in vivo* of striatal sites tuned to low frequencies suggested that the pattern of potentiation might be spatially organized within the striatum. We therefore developed an *in vitro* brain slice preparation to examine this possibility. We first characterized the tonotopic organization of the auditory corticostriatal projection by injecting adeno-associated viruses encoding either red or green fluorescent proteins at two different positions along the auditory cortical tonotopic axis. Cortical axons terminated in the striatum in distinct bands, with cortical projections tuned to high frequency sounds terminating more laterally in the auditory striatum and projections tuned to low-frequency sounds more medially (Fig. 4a and Extended Data Fig. 5). We next developed a protocol to assess the gradient of corticostriatal potentiation along the tonotopic axis, by recording ChR2-LFPs in coronal slices that preserve striatal tonotopy (Fig. 4b, see Methods). These recordings targeted left striatum, contralateral to the reward direction associated with low-frequency stimuli (LowRight; $n = 7$ rats). For consistency across experiments, we used only a single slice from each animal, selected on the basis of striatal and hippocampal landmarks (see Methods). ChR2-LFPs in these slices showed a stereotyped waveform similar to that observed *in vivo*, and pharmacological dissection confirmed that the late component of the response was mediated mainly by AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid)-type glutamate receptors (Fig. 4c). Simultaneous extracellular and intracellular recording indicated that the ChR2-LFP was a faithful measure of synaptic strength (Fig. 4d, e, top). As expected, the normalization corrected for changes in ChR2-LFP induced by recruiting more presynaptic fibres (Fig. 4e, bottom), but did not obscure true changes in synaptic strength induced by changes in release probability (Fig. 4d, bottom).

To measure gradients in synaptic strength along the tonotopic axis induced by training, in each slice we recorded the ChR2-LFP at between

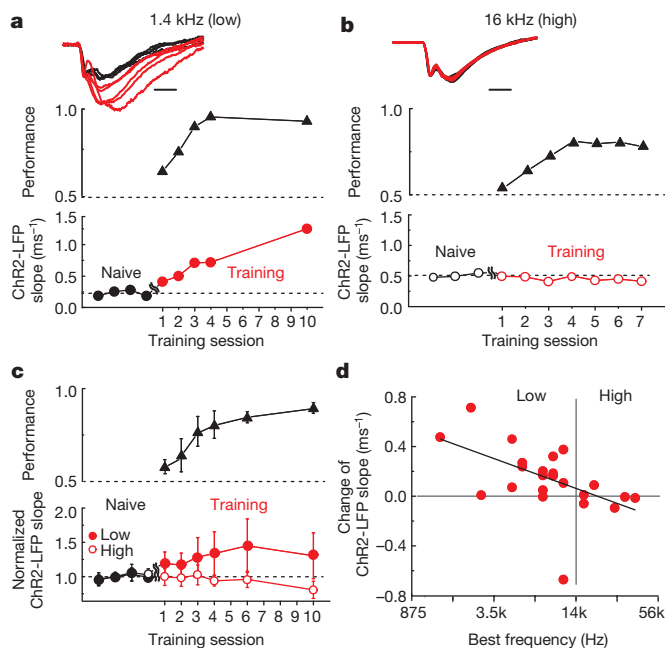


Figure 2 | Frequency-selective potentiation of corticostriatal ChR2-LFP slope during learning. **a, b**, ChR2-LFP (LFP slope: see Methods) before (black) and during (red) training at example sites tuned to low (**a**) and high frequency (**b**). Session 1 is defined as the first session in which the animal performed the full task (see Methods). Scale bars, 2 ms. **c**, Population average of normalized (see Methods) ChR2-LFP slope during learning for sites tuned to low (<14 kHz, $n = 16$ sites, filled circles) and high (>14 kHz, $n = 6$ sites, open circles) frequencies. Mean \pm s.e.m. **d**, Potentiation is restricted to sites tuned to low (<14 kHz) frequencies (23 recording sites from eight rats; least squares regression of potentiation against frequency, $P = 0.011$).

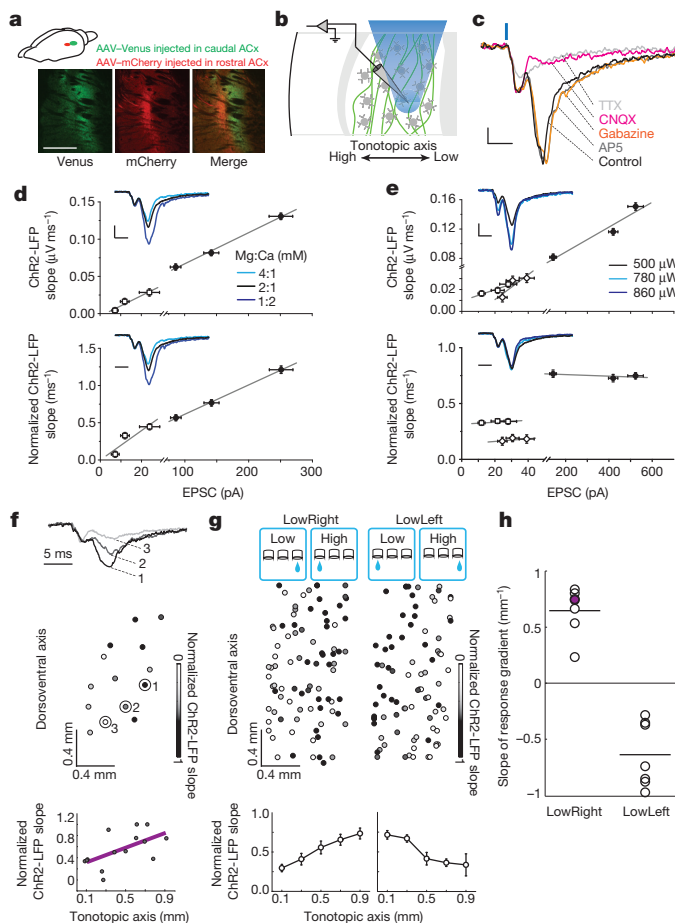


Figure 4 | Gradient of corticostriatal ChR2-LFP slopes encodes the association between stimulus and action. **a**, Tonotopy of projections from auditory cortex to striatum. Scale bar, 1 mm. **b**, Recording method. Light spot (blue) activates a subset of ChR2-expressing corticostriatal axons (green) near recording site. **c**, Pharmacological dissection of ChR2-LFP in a striatal slice. Scale bars, 50 μ V and 5 ms. **d**, Paired *in vitro* excitatory postsynaptic current and LFP at different external divalent ion concentrations. Slopes of LFP measured from raw traces (upper row) and normalized traces (lower row) changed linearly with excitatory postsynaptic current amplitudes ($R^2 = 0.96$ and 0.99 for circles, $R^2 = 0.94$ and 0.81 for squares in upper and lower row respectively; grey lines in right panels are linear regression fits for each recording pair). **e**, Paired recording at different light levels. Slopes of the LFPs measured from raw traces (upper row) changed monotonically with excitatory postsynaptic current amplitudes ($R^2 = 0.94$ for filled circles, $R^2 = 0.91$ for squares and 0.77 for diamonds). Slopes of the normalized LFPs remain constant (lower row). Scale bars for **d** and **e**, 100 μ V and 5 ms. **f**, Normalized ChR2-LFP recorded at many sites within a striatal slice. Sample waveforms (1–3) shown above. ChR2-LFP slope increases with position along tonotopic axis (lower panel). **g**, Population data for LowRight ($n = 7$ rats) and LowLeft ($n = 7$ rats). Error bars, s.e.m. **h**, Gradient correctly identifies learned association in 14 out of 14 individual rats (binomial test $P = 0.00006$). Slope of example shown in **f** is indicated in purple. Bars, mean values.

8 and 16 sites (12.1 ± 2.1). Naive rats showed no systematic difference in the strength of cortical input along the striatal tonotopic axis (Extended Data Fig. 6). In contrast, in rats trained to associate low frequencies with rightward choices, the evoked corticostriatal response was strongest at medial (low frequency) sites and decreased laterally (Fig. 4f, g). This confirmed our observations *in vivo* and was consistent with the association to contralateral rewards. Thus, the degree of corticostriatal synaptic potentiation induced by learning depended systematically on the position along the striatal tonotopic axis.

If the gradient of potentiation along the striatal tonotopic axis encodes the association between frequency and choice direction, then animals

trained to make the opposite association should have a gradient of opposite sign. To test this we trained a new cohort of animals to associate low frequencies with leftward choices (LowLeft; $n = 7$ rats). As predicted, the gradient in these animals was of similar magnitude but opposite in sign (Fig. 4g). There was no difference between these two training groups in ChR2-LFP across the orthogonal (dorsoventral) axis ($P = 0.22$, paired t -test; Extended Data Fig. 7). Thus the spatial gradient of corticostriatal potentiation induced by learning along the tonotopic axis depends on the training contingencies to which the animal is subjected.

Finally, we wondered whether the direction of the stimulus–response association could be inferred on the basis of the sign of the ChR2-LFP gradient in individual animals. Remarkably, the training history (LowRight versus LowLeft) of every rat (14 out of 14) could be correctly inferred from the sign of gradient in a single slice (binomial test $P = 0.00006$, Fig. 4h). The correlation between synaptic strength and tonotopic position reached statistical significance ($P < 0.05$) in 6 out of 14 slices. Thus post-mortem study of corticostriatal efficacy can reliably reveal the training history of individual subjects.

Our results suggest a simple model of how the specific pattern of corticostriatal potentiation we observed might mediate task acquisition. In the LowRight task, training selectively potentiated corticostriatal synapses tuned to low frequencies between the left auditory cortex and the left auditory striatum (Extended Data Fig. 8). Thus in behaving animals, low-frequency tones would trigger stronger activation in the left auditory striatum and direct the animal to the right (contralateral) response port, possibly through the action of direct pathway medium spiny neurons¹⁰ that project ipsilaterally to the left substantia nigra pars reticulata and in turn to the superior colliculus¹⁶. On the other hand, in LowLeft-trained animals, potentiation would cause the same stimulus to trigger stronger activation in right auditory striatum and direct the animal to the left response port. Although this model ignores much of the complexity of striatal circuitry, it offers a simple framework for understanding our results.

Previous work has demonstrated synaptic^{17–20} or receptive field^{21–26} changes induced by learning. Our results identify a locus of synaptic plasticity during the acquisition of a sensory discrimination task. We focused on auditory frequency discrimination, which allowed us to exploit the spatial organization of auditory corticostriatal connections to relate the tuning of cortical neurons to plasticity. Training selectively enhanced the strength of cortical inputs to establish an orderly gradient of corticostriatal synaptic strength across the striatal tonotopic axis.

The strengthening of a subset of connections, selected from a rich sensory representation, is reminiscent of several powerful models of learning^{27–29}. In these models, even difficult nonlinear classification can be achieved by combining a high-dimensional representation of the stimulus—such as is found in the auditory cortex³⁰—through simple reinforcement learning rules. We speculate that selective strengthening of appropriate corticostriatal synapses would allow animals to categorize a wide range of sensory stimuli—even those that are not mapped topographically in the striatum—and may reflect a general mechanism through which sensory representations guide the selection of motor responses.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 September 2014; accepted 9 January 2015.

Published online 2 March 2015.

1. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013).
2. Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346**, 174–177 (1990).
3. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
4. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nature Neurosci.* **6**, 1224–1229 (2003).

5. Felsen, G. & Mainen, Z. F. Neural substrates of sensory-guided locomotor decisions in the rat superior colliculus. *Neuron* **60**, 137–148 (2008).
6. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).
7. Raposo, D., Sheppard, J. P., Schrater, P. R. & Churchland, A. K. Multisensory decision-making in rats and humans. *J. Neurosci.* **32**, 3726–3735 (2012).
8. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
9. Thompson, J. A. & Felsen, G. Activity in mouse pedunculopontine tegmental nucleus reflects action and outcome in a decision-making task. *J. Neurophysiol.* **110**, 2817–2829 (2013).
10. Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nature Neurosci.* **15**, 1281–1289 (2012).
11. Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z. & Graybiel, A. M. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158–1161 (2005).
12. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).
13. Reynolds, J. N., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).
14. Yin, H. H. *et al.* Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neurosci.* **12**, 333–341 (2009).
15. Malenka, R. C. & Kocsis, J. D. Presynaptic actions of carbachol and adenosine on corticostriatal synaptic transmission studied *in vitro*. *J. Neurosci.* **8**, 3750–3756 (1988).
16. Hikosaka, O., Takikawa, Y. & Kawagoe, R. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* **80**, 953–978 (2000).
17. Carew, T. J., Walters, E. T. & Kandel, E. R. Associative learning in *Aplysia*: cellular correlates supporting a conditioned fear hypothesis. *Science* **211**, 501–504 (1981).
18. Rioult-Pedotti, M. S., Friedman, D. & Donoghue, J. P. Learning-induced LTP in neocortex. *Science* **290**, 533–536 (2000).
19. Rumpel, S., LeDoux, J., Zador, A. & Malinow, R. Postsynaptic receptor trafficking underlying a form of associative learning. *Science* **308**, 83–88 (2005).
20. Whitlock, J. R., Heynen, A. J., Shuler, M. G. & Bear, M. F. Learning induces long-term potentiation in the hippocampus. *Science* **313**, 1093–1097 (2006).
21. Finnerty, G. T., Roberts, L. S. & Connors, B. W. Sensory experience modifies the short-term dynamics of neocortical synapses. *Nature* **400**, 367–371 (1999).
22. Trachtenberg, J. T. *et al.* Long-term *in vivo* imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* **420**, 788–794 (2002).
23. Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. A synaptic memory trace for cortical receptive field plasticity. *Nature* **450**, 425–429 (2007).
24. Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T. & Hubener, M. Experience leaves a lasting structural trace in cortical circuits. *Nature* **457**, 313–317 (2009).
25. Fritz, J., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neurosci.* **6**, 1216–1223 (2003).
26. Edeline, J. M. & Weinberger, N. M. Receptive field plasticity in the auditory cortex during frequency discrimination training: selective retuning independent of task difficulty. *Behav. Neurosci.* **107**, 82–103 (1993).
27. Maass, W., Natschlager, T. & Markram, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
28. Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
29. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
30. Hromádka, T., Deweese, M. R. & Zador, A. M. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol.* **6**, e16 (2008).

Acknowledgements We thank B. Burbach for technical help, R. Eifert for mechanical material support, and J. Cohen for training the rats. AAV-CAGGS-ChR2-Venus was provided by K. Svoboda. We thank U. Livneh and A. Reid for discussions. This work was supported by grants (R01DC012565 and R01NS088649) from the National Institutes of Health and the Swartz Foundation (A.M.Z.).

Author Contributions Q.X., P.Z., and A.M.Z. designed the experiments; Q.X. performed the experiments; Q.X., P.Z., and A.M.Z. analysed the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.Z. (zador@cshl.edu).

METHODS

No statistical methods were used to predetermine sample size.

Animals and viruses. Animal procedures were approved by the Cold Spring Harbour Laboratory Animal Care and Use Committee and performed in accordance with National Institutes of Health standards. AAV-CAGGS-ChR2-Venus serotype 2/9 was packaged by the University of Pennsylvania Vector Core.

Long Evans male rats (Taconic Farm) were anaesthetized with a mixture of ketamine (50 mg per kg) and medetomidine (0.2 mg per kg), and injected with virus at 3–4 weeks old in the left auditory cortex. To cover most of the area and layers of the primary auditory cortex, three or four injections were made perpendicularly to the brain surface at 1, 2, and 3 mm caudal to the temporoparietal suture, and 1 mm from the ventral edge. Each injection was made at three depths (500, 750, and 1,000 μm), expelling approximately 200 nl of virus at each depth.

Behavioural training. Rats were placed on a water deprivation schedule and trained to perform an auditory two-alternative forced-choice task in a single-walled sound-attenuating training chamber as described previously¹. Briefly, freely moving rats were trained to initiate a trial by poking into the centre port of a three-port operant chamber, which triggered the presentation of a stimulus. Subjects then selected the left or right goal port. Correct responses were rewarded with water. The cloud-of-tones stimulus consisted of a stream of 30-ms overlapping pure tones presented at 100 Hz. The stream of tones continued until the rat withdrew from the centre port. Eighteen possible tone frequencies were logarithmically spaced from 5 to 40 kHz. For each trial either the low stimulus (5–10 kHz) or high stimulus (20–40 kHz) was selected as the target stimulus, and the rats were trained to report low or high by choosing the correct side of port for water reward. In the LowLeft task, the rats were required to go to the left goal port for water reward when the low stimulus was presented, and to the right goal port when high stimulus was presented. In the LowRight task, the rats were required to go to the right goal port when the low stimulus was presented, and to the left goal port when the high stimulus was presented.

The pre-stimulus delay was drawn from exponential distribution with a mean of 300 ms. Early withdrawal from the centre port before the onset of stimuli terminated the trial and a new trial was started. To complete a trial after exiting the centre port, the animals were allowed up to 3 s to select a reward port. Typically, they made their choice within 300–700 ms. Error trials, where the rats reported to the wrong goal port after the presentation of the stimulus, were penalized with a 4 s time-out.

The intensity of individual tones was constant during each trial. To discourage rats from using loudness differences in discrimination, tone intensity was randomly selected on each trial from a uniform distribution between 45 and 75 dB (sound pressure level) during training.

Implanted rats were water deprived and given free water for 1 h every day before ChR2-LFP recording. These sessions were used to record baseline ChR2-LFP responses and defined as naive sessions. Once a stable baseline was achieved, we began training subjects to perform the cloud-of-tones task. To introduce the subjects to the task structure, they were first trained ('direct mode') to poke at the centre port, which triggered the presentation of a stimulus and elicited water delivery from the corresponding goal port. Direct mode training was continued until a subject completed at least 150 trials in a single session (usually the first or second session). In subsequent sessions, defined as 'session 1' in Figs 2 and 3, the animal was trained on the 'full task', in which water was delivered only if the subject poked the correct goal port. For control subjects used in Fig. 3, animals were trained in the direct mode with visual stimuli before introducing the full visual task, and then trained on the full auditory task.

Tetrode recording and optogenetics. Custom tetrode and optic fibre arrays were assembled as described previously¹. Each array carried six individually movable microdrives. Each microdrive consisted of one tetrode (4 polyimide-coated nichrome; wire diameter 12.7 μm ; Kanthal Palm Coast) twisted together and gold-plated to an impedance of 0.3–0.5 M Ω at 1 kHz) and one optic fibre (62.5 μm diameter with a 50- μm core; Polymicro Technologies). The tetrode and fibre on the same microdrive were glued together, with the tips approximately 100 μm from each other.

To implant the tetrode/fibre array, rats were anaesthetized with a mixture of ketamine (50 mg per kg) and medetomidine (0.2 mg per kg) and placed in a stereotaxic apparatus. A craniotomy was made above the target area (2.5–3.5 mm from the bregma and 4–6 mm lateral from the midline). All rats were implanted in left hemisphere. The array was fixed in place with dental acrylic, and the tetrodes were lowered down to auditory striatum (3–5 mm from pia).

Electrical signals in the auditory striatum were recorded using a Neuralynx Cheetah 32-channel system and Cheetah data acquisition software. For action potential recording, signals were filtered 600–6,000 Hz. For LFP recording, signals were filtered 10–9,000 Hz. The rise time of the ChR2-LFP is relatively fast, so to preserve its dynamics we sought to stay as close to the raw data as possible. We re-analysed the data in Fig. 2a using two other choices of offline filter (median, 660-ms window and Butterworth lowpass 1–800 Hz). As shown in Extended Data Fig. 9a, b, although filtering

does affect the details of the ChR2-LFP shape (especially the earliest component), the results are qualitatively unchanged (Extended Data Fig. 9c).

To determine the preferred sound frequency of recording sites, pure tones spanning 1–64 kHz were presented to rats before the start of behavioural training in a soundproof chamber for 100 ms every 2 s, in a random order at 30, 50 or 70 dB (sound pressure level)¹. The multi-unit baseline-subtracted firing rate in a window 5–55 ms after sound onset was compared with that in a window 0–50 ms before sound onset; only sites that significantly responded to sound were included. Firing rates in the window 5–55 ms after sound onset were computed for each frequency at 70 dB, and the peak of the resulting tuning curve was selected as the preferred frequency.

For ChR2-LFP recording, 473 nm laser light was delivered through an FC/PC patch cord using a FibrePort Collimator (Thor Labs) to each implanted fibre individually. ChR2-LFP was recorded immediately after each training session. The laser power out of the patch cord was measured and adjusted to elicit an LFP with clear early and delayed components at each recording site (1–10 mW). For individual recording sites, laser power was adjusted slightly between days to maintain the early, presynaptic component of the LFP response at a consistent level. Each light pulse was 100 μs in duration, presented at 1 Hz, and each recording was an average of approximately 100 trials.

In vivo pharmacology. To dissect the components of ChR2-LFP *in vivo*, rats were anaesthetized and placed in a stereotaxic apparatus. A single tetrode/fibre bundle was placed on a motorized manipulator (Sutter Instrument Company) and the tips of tetrode/fibre were guided to the auditory striatum. Glass pipettes were used to deliver chemicals into the auditory striatum. The pipettes filled with the desired chemicals were carefully moved to penetrate through the cortex and were placed with the tips close to the auditory striatum. Air pressure was slowly applied to inject the chemicals into tissue through a syringe that was connected to the pipette.

Slice recording. Virus-injected and trained rats were anaesthetized and decapitated, and the brains were transferred to a chilled cutting solution composed of (in mM) 110 choline chloride, 25 NaHCO₃, 25 D-glucose, 11.6 sodium ascorbate, 7 MgCl₂, 3.1 sodium pyruvate, 2.5 KCl, 1.25 NaH₂PO₄, and 0.5 CaCl₂. Coronal slices (350 μm) were cut and transferred to artificial cerebrospinal fluid containing (in mM) 127 NaCl, 25 NaHCO₃, 25 D-glucose, 2.5 KCl, 4 MgCl₂, 1 CaCl₂, and 1.25 NaH₂PO₄, aerated with 95% O₂ 5% CO₂.

To ensure maximal alignment across animals, only a single slice (350 μm thickness, between 2.5 and 2.9 mm from the Bregma) per animal was used. Slices were incubated at 34 °C for 15–20 min and then kept at room temperature (22 °C) during the experiments. LFPs were recorded using Axopatch 200B amplifiers (Axons Instruments, Molecular Devices).

We delivered light pulses through a light guide microscope illumination system (Lumen Dynamics) modified to accept a blue laser (473 nm, Lasermate Group) in place of the lamp. The laser beam was focused onto the sample through the $\times 60$ objective during recordings, with an illumination field of 350 μm diameter. Each light pulse was 500 μs at 1 Hz, and each recording was an average of approximately ten trials. Laser power was adjusted from site to site to maintain a similar level of axonal stimulation as judged by the amplitude of the early, presynaptic component of the ChR2-LFP response. To minimize the contribution of rundown on the estimation of the plasticity gradient within the striatal slice, recording locations were selected randomly for each slice.

For drug application, 50 μM AP5, 5 μM Gabazine, 50 μM CNQX, and 0.5 μM TTX were sequentially delivered through a perfusion system.

Data analysis. All data were analysed in MATLAB.

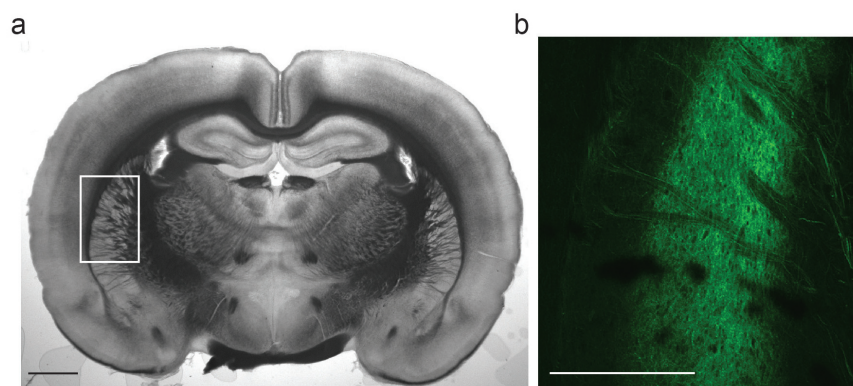
Behaviour analysis only included completed trials. The percentage of correct trials for each animal in each session was computed using the last 200 trials of that session, unless the number of trials was fewer than 300, in which case only the last 100 trials were used.

Each measurement of *in vivo* ChR2-LFP was from a trace obtained by averaging across 70–100 trials (the slope measured from the averaged trace was not different from the averaged slope of single traces; Extended Data Fig. 10). Each average trace was normalized to the peak of the first component (around the time window between 0.5 and 1.2 ms after light stimulation; Extended Data Fig. 2a), and the LFP slope was estimated by a linear regression fit of the rising phase of the second component (Extended Data Fig. 2b; the same time window was used for each recording site across sessions, the time window from site to site varied and was adjusted by eye for each site, ranging from 1.6 to 5 ms after light stimulation). The ChR2-LFP slope for each recording site across sessions was used as a measure of synaptic strength in subsequent analyses (Figs 2a, b and 3b). The absolute change of the ChR2-LFP slope was used in Figs 2d and 3d. In population analyses, normalized synaptic strength for each site was obtained by dividing the LFP slope values from all sessions by the mean of the ChR2-LFP slope values from naive sessions (Figs 2c and 3c).

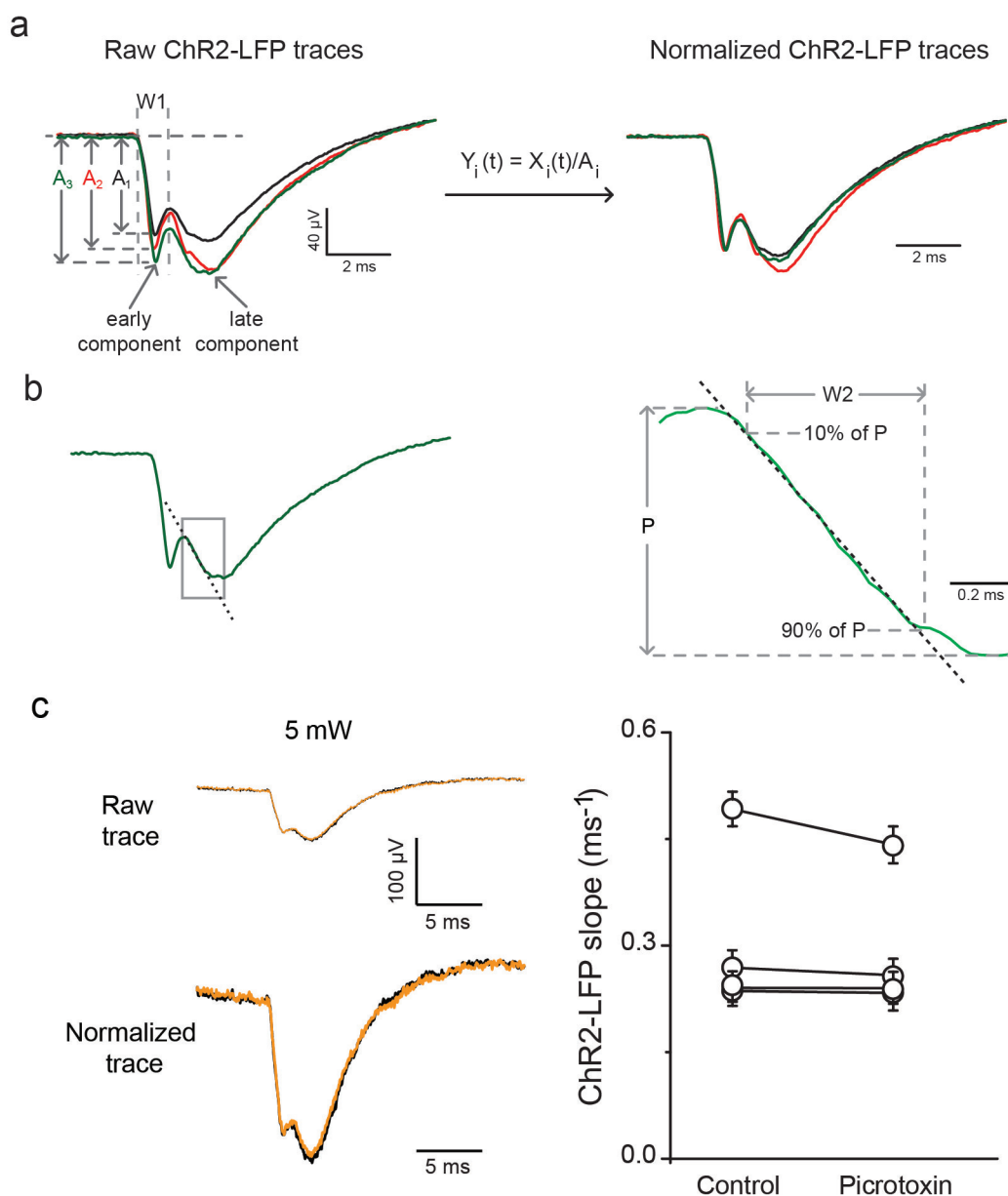
For quantification of the ChR2-LFP in slice recording, the ChR2-LFP slope at each recording site was obtained in a manner similar to that *in vivo*: each averaged trace was normalized to the peak of the first component (around the time window

between 1.5 and 4 ms after light stimulation), and the ChR2-LFP slope was estimated by a linear regression fit of the rising phase of the second component (around the time window between 5 and 8 ms after light stimulation, adjusted by eye for each slice). For each slice, the ChR2-LFP slopes across sites were re-scaled from 0 to 1,

with the smallest ChR2-LFP set to zero and the largest to 1. All recorded brain slices were aligned to a consensus brain slice. The positions of recording sites were measured from the aligned brain slices. Data from all the slices were pooled together for plotting the summary plot and for the quantitative analysis.

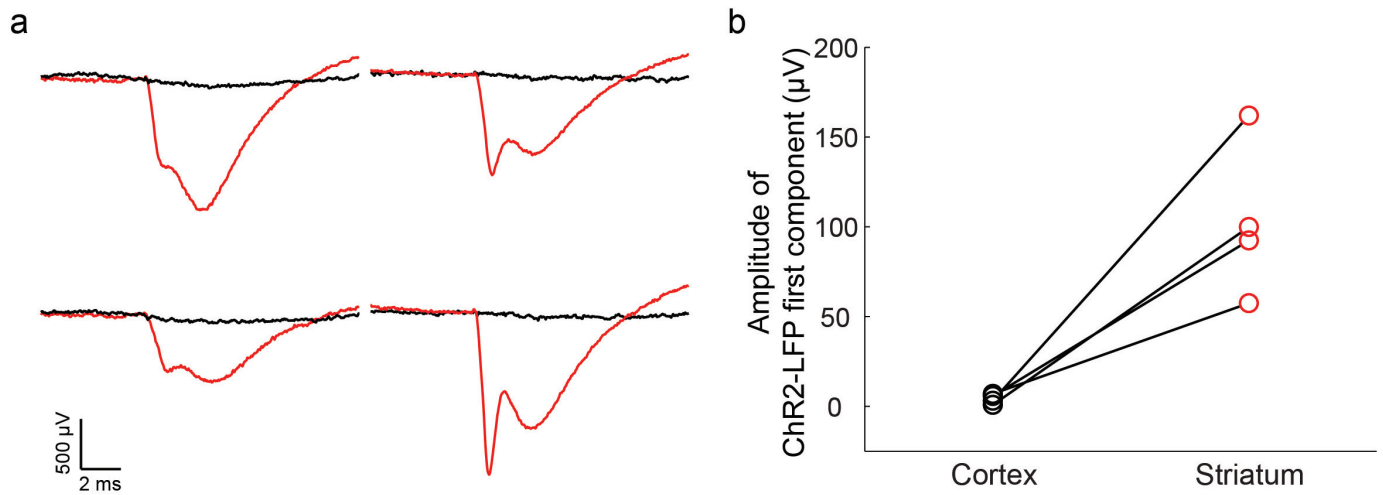


Extended Data Figure 1 | Corticostriatal projections from auditory cortex to striatum. **a**, Coronal view for the location in the striatum that receives auditory cortical inputs. **b**, Confocal image of auditory cortical axon terminals expressing ChR2-Venus in the striatum. Scale bars, 2 mm.



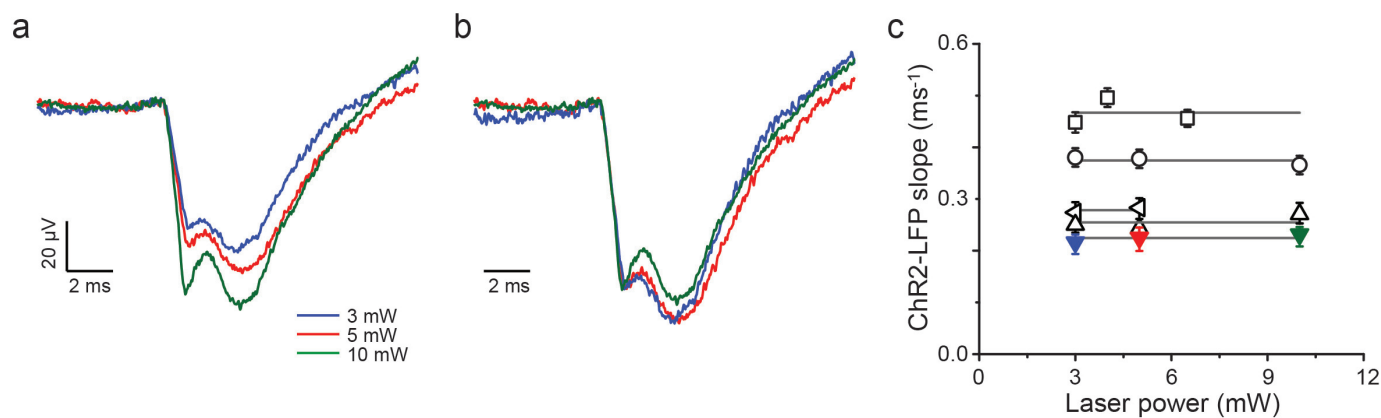
Extended Data Figure 2 | Slope measurement for ChR2-LFP and GABAergic (γ -aminobutyric-acid-mediated) synaptic transmission does not contribute to the ChR2-LFP slope *in vivo*. **a**, Raw ChR2-LFP traces (left) were normalized to the amplitude of their corresponding early component (A_i). The normalization factor A_i was determined as the peak of the raw trace in the time window ($W1$) between 0.5 and 1.2 ms after light stimulation onset. **b**, The rising phase of the late component of ChR2-LFP (in a time window $W2$ defined by rise from 10% to 90% of the peak P) was fitted linearly, and the slope

of the fit was used for the quantification of ChR2-LFP. **c**, Left: ChR2-LFP before (black traces) and after (orange traces) picrotoxin application (20 mM, 5 μl). Raw traces are averaged traces from 60–80 trials at each condition (upper row). Normalized traces are raw traces normalized to their peaks of first components (as illustrated in **a**). Right: slopes measured from normalized traces in control and picrotoxin conditions for each recording before and after picrotoxin application ($P = 0.8$, paired signed-rank test). Data are presented as mean \pm s.e.m.



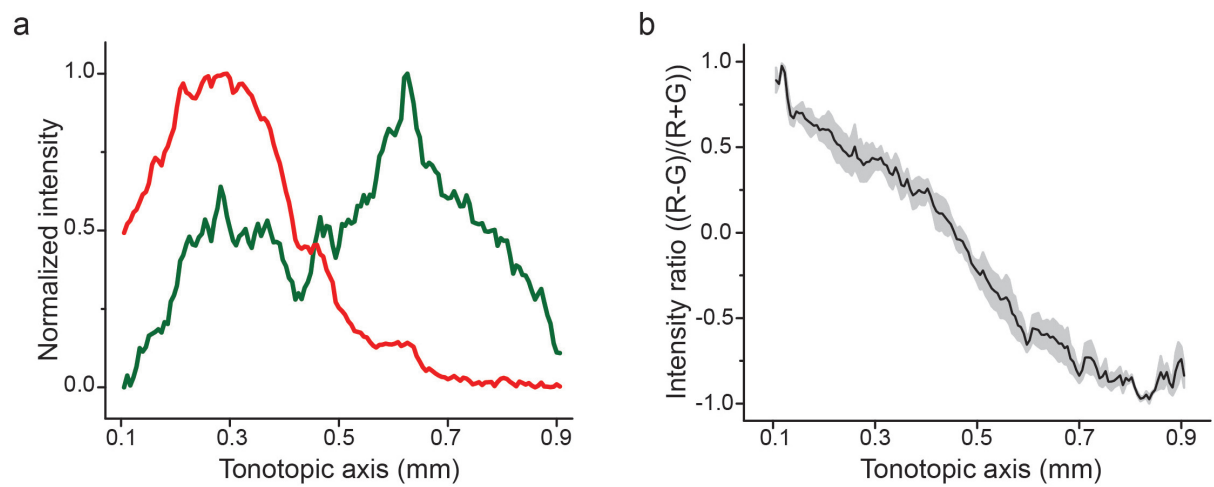
Extended Data Figure 3 | ChR2-LFP depends on the presence of ChR2-expressing axons. To rule out the possibility that the TTX-insensitive component of the light-evoked response resulted from a photoelectric or other artefact, rather than from ChR2-evoked currents, we assessed light-evoked responses in brain regions that did not express ChR2. **a**, Four independent recordings in the auditory striatum (red traces) which receives auditory cortical

input (ChR2-expressing axons are present), and the overlying somatosensory cortex (black traces) which lacks auditory cortical input (ChR2-expressing axons are absent). Each pair of recordings is from the same tetrode/fibre bundle. The recordings indicate that the light artefact was negligible under our conditions. **b**, Comparison of the first component amplitude from each recording pair.

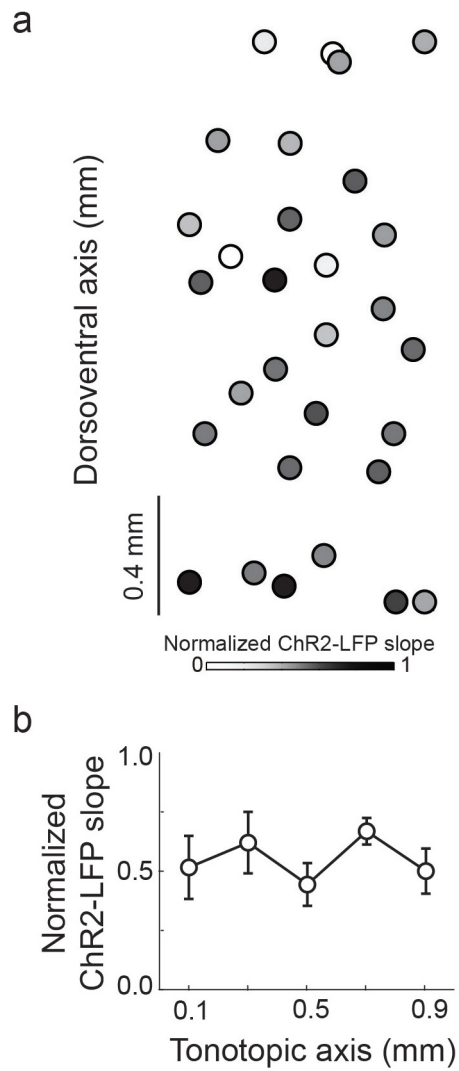


Extended Data Figure 4 | Normalization procedure corrects for variation in light power *in vivo* (for *in vitro* data see Fig. 4d, e). **a**, Example of ChR2-LFP recorded at different light levels. **b**, Normalized ChR2-LFP, the same example as in **a**. **c**, Slopes from five example recordings across 1–10 mW light level range (coloured symbols are examples shown in **a** and **b**). Grey lines are

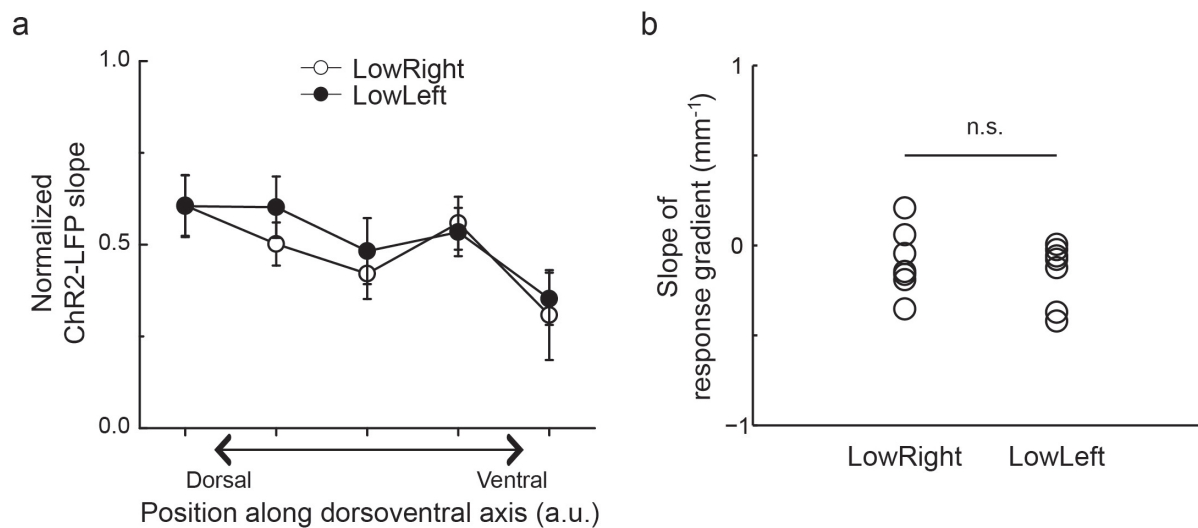
drawn from the mean values of each group. Together with the data shown in Fig. 4e, the normalization procedure thus minimizes fluctuations in the response arising from artefactual changes in the number of recruited fibres, but preserves changes arising from actual increases or decreases in synaptic efficacy.



Extended Data Figure 5 | Quantification of corticostriatal projection topography. **a**, Normalized red and green fluorescence intensities measured across the tonotopic axis from the image shown in Fig. 4a. **b**, Mean red:green intensity ratio across the tonotopic axis ($n = 3$ sections from two rats). Shading, s.e.m.

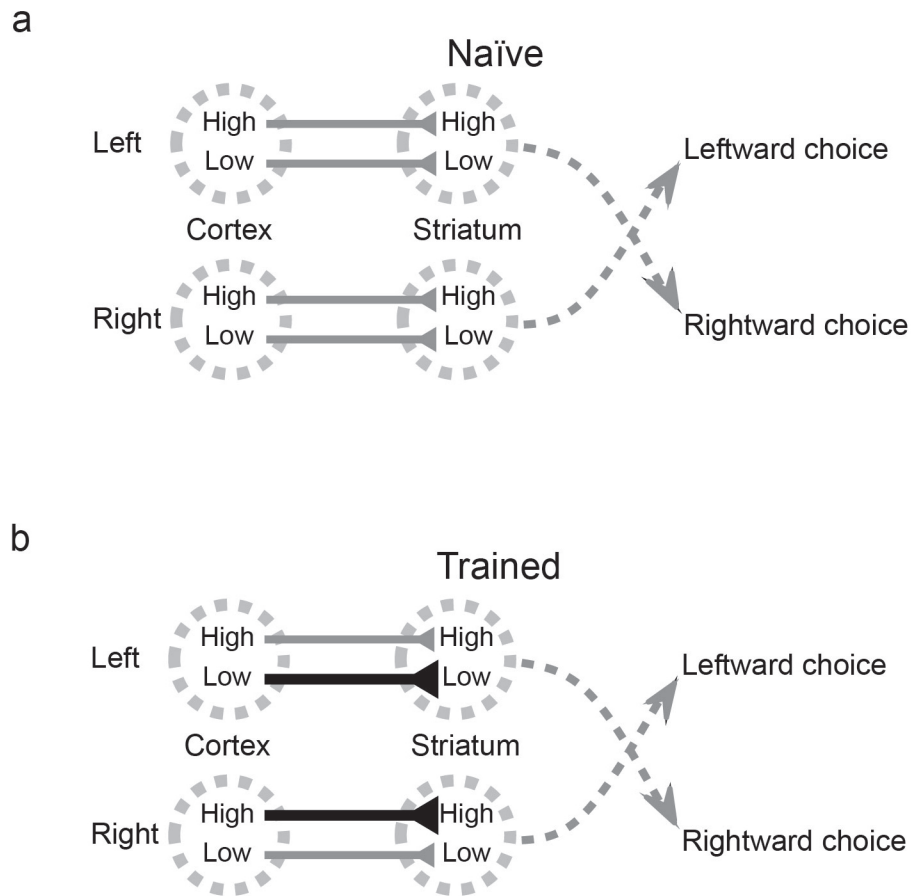


Extended Data Figure 6 | ChR2-LFP slope does not vary systematically across the tonotopic axis in naive rats. a, ChR2-LFP slope map from three striatal slices ($n = 3$ rats). **b,** Quantification of the ChR2-LFP slope across the tonotopic axis. Data are mean \pm s.e.m.



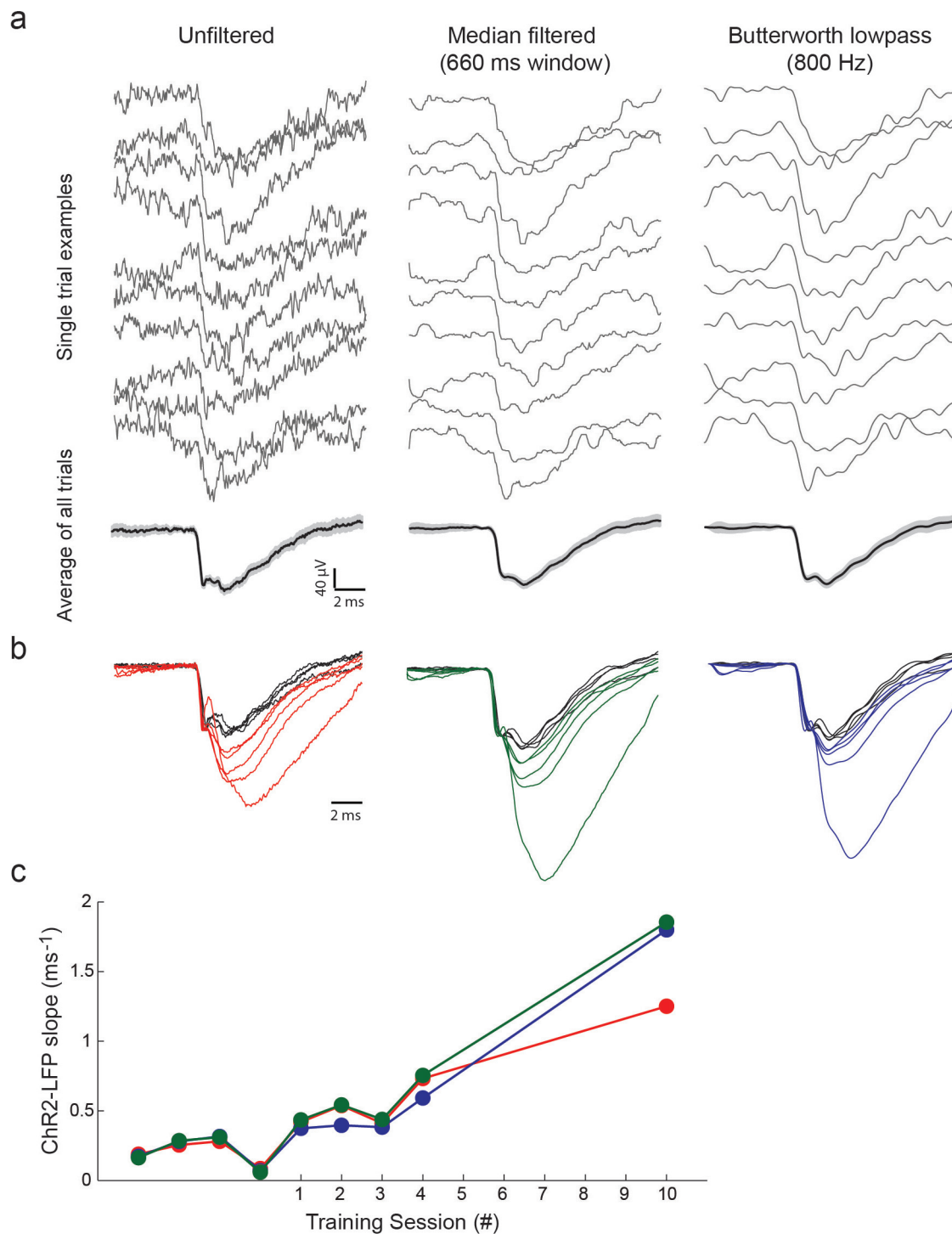
Extended Data Figure 7 | Gradient of ChR2-LFP across the dorsoventral (non-tonotopic) axis showed no difference between the two training groups.
a, Averaged ChR2-LFP slopes with position along the tonotopic axis for

LowRight and LowLeft groups (seven rats from each group). **b**, Individual gradients of ChR2-LFP across the dorsoventral axis from LowRight and LowLeft groups ($P = 0.22$, paired t -test).



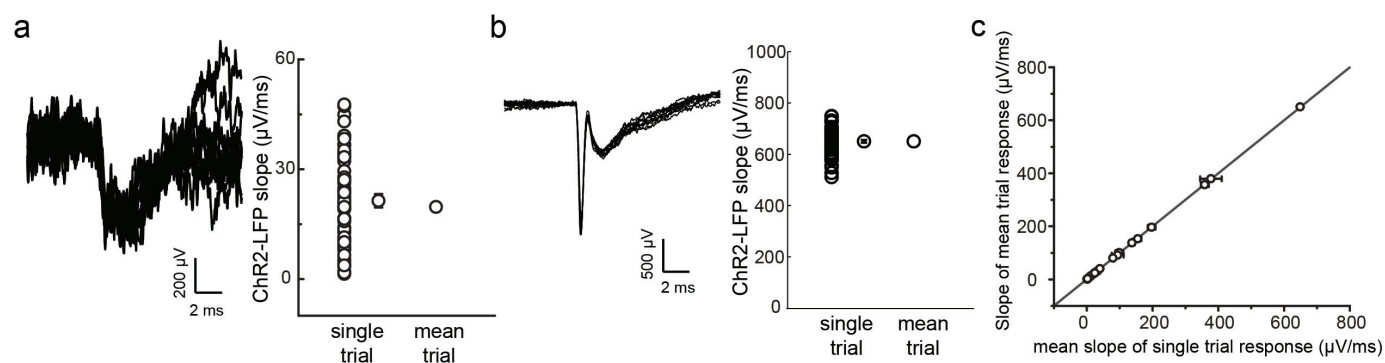
Extended Data Figure 8 | Model showing how corticostriatal potentiation could mediate task acquisition. **a**, In the naïve rat, the strength of corticostriatal connections does not depend on their frequency preference. **b**, Training to associate low stimuli with rightward choices and high stimuli

with leftward choices (LowRight) selectively potentiates corticostriatal synapses tuned to low frequencies in the left hemisphere and corticostriatal synapses tuned to high frequencies in the right hemisphere. Thus, in the trained rat, low stimuli drive rightward choices and high stimuli drive leftward choices.



Extended Data Figure 9 | To exclude the possibility that spiking responses affected the ChR2-LFP measurement, we analysed the data after median or lowpass filtering. **a**, Single trial (upper rows) and average (bottom row) examples of unfiltered, median filtered, and Butterworth lowpass filtered

responses. Average traces are presented as mean values (black traces) with 95% confidence intervals (grey shading). **b**, ChR2-LFP examples in Fig. 2a with different filter settings. **c**, ChR2-LFP measurements from examples shown in Fig. 2a at different filter settings.



Extended Data Figure 10 | Changes in ChR2-LFP could result from variation in response timing precision. To rule out this possibility we compared slopes measured from single trial and mean responses. **a**, Single trial responses (left) and slopes measured from individual trials and mean response

(right) at a weakly light-responsive site. **b**, An example robustly responsive site. **c**, Comparison of mean slopes from single trial responses and slopes quantified from mean responses.

Early reprogramming regulators identified by prospective isolation and mass cytometry

Ernesto Lujan^{1,2,3*}, Eli R. Zunder^{4*}, Yi Han Ng^{1,3,5}, Isabel N. Goronzy^{1,3}, Garry P. Nolan⁴ & Marius Wernig^{1,3}

In the context of most induced pluripotent stem (iPS) cell reprogramming methods, heterogeneous populations of non-productive and staggered productive intermediates arise at different reprogramming time points^{1–11}. Despite recent reports claiming substantially increased reprogramming efficiencies using genetically modified donor cells^{12,13}, prospectively isolating distinct reprogramming intermediates remains an important goal to decipher reprogramming mechanisms. Previous attempts to identify surface markers of intermediate cell populations were based on the assumption that, during reprogramming, cells progressively lose donor cell identity and gradually acquire iPS cell properties^{1,2,7,8,10}. Here we report that iPS cell and epithelial markers, such as SSEA1 and EpCAM, respectively, are not predictive of reprogramming during early phases. Instead, in a systematic functional surface marker screen, we find that early reprogramming-prone cells express a unique set of surface markers, including CD73, CD49d and CD200, that are absent in both fibroblasts and iPS cells. Single-cell mass cytometry and prospective isolation show that these distinct intermediates are transient and bridge the gap between donor cell silencing and pluripotency marker acquisition during the early, presumably stochastic, reprogramming phase². Expression profiling reveals early upregulation of the transcriptional regulators *Nr0b1* and *Etv5* in this reprogramming state, preceding activation of key pluripotency regulators such as *Rex1* (also known as *Zfp42*), *Dppa2*, *Nanog* and *Sox2*. Both factors are required for the generation of the early intermediate state and fully reprogrammed iPS cells, and thus represent some of the earliest known regulators of iPS cell induction. Our study deconvolutes the first steps in a hierarchical series of events that lead to pluripotency acquisition.

Reprogramming somatic cells to a pluripotent state by forced transcription factor expression is typically an inefficient process involving heterogeneous populations that impede molecular analysis of productive reprogramming^{1–11,14}. Previous studies have shown that reprogramming is a multi-stage process involving presumably early stochastic and late deterministic phases^{2,3,5}. Progress has been made characterizing intermediates of the late phase given the appearance of well-known pluripotency markers at that time^{1,2,7,8,10,15}. In contrast, not much is known about the early stochastic phase except the consistent observation that downregulation of donor cell markers is an early feature of successful reprogramming^{1,2,7,8,10,15–20}.

To identify surface markers of early reprogramming stages, we screened 176 antibodies on cells representing three stages of the reprogramming process: (1) mouse embryonic fibroblasts (MEFs), (2) a previously characterized partially reprogrammed cell (PRC) line^{18,20,21} and (3) embryonic stem cells (ESCs). We identified 21 markers enriched or shared between these cell types and characterized their co-expression by single-cell mass cytometry using spanning-tree progression analysis of density-normalized events (SPADE), which groups similar cells into a defined number of clusters^{22,23} (Fig. 1a, b and Extended Data Fig. 1). Next, we characterized their expression by mass cytometry during *Oct4*-, *Sox2*-,

Klf4- and *c-Myc*-driven MEF reprogramming. By day 3, downregulation of the fibroblast expression program was evident (Fig. 1c and Extended Data Figs 2 and 3). At day 6, major branches were delineated by the PRC marker CD73 and ESC markers CD54, CD326 and SSEA1. Little co-expression was observed between these markers, suggesting several intermediates arise during early reprogramming or early expression of some of these markers may not be indicative of productive reprogramming. By day 9, CD326 and SSEA1 expression converged in a subpopulation and persisted on days 12 and 16 (Extended Data Fig. 2b). These clusters were heterogeneous for CD73, suggesting they may be derivatives of separate populations or a CD73^{high} subpopulation becomes CD326^{high}, SSEA1^{high}. Over our time course, the ESC marker CD54 largely localized to fibroblast branches and did not cluster with CD326^{high} and SSEA1^{high} clusters, suggesting CD54 expression in pluripotent cells is a late event⁷.

Hypothesizing that cells destined to successfully reprogram acquire surface markers in a stepwise, non-stochastic manner during early reprogramming, we assessed reprogramming efficiencies for cells with high or low expression of the above-characterized surface markers at early time points (Fig. 2a–c and Extended Data Fig. 4a). As mass culture experiments can be misleading because a single proliferative population can seed multiple secondary colonies, we conducted 96-well assays to assess unique reprogramming events. On the basis of the current literature it would be expected that reprogramming cells would be enriched by (1) low levels of fibroblast markers and (2) high levels of ESC markers independent of the time and state of reprogramming^{2,7,8,10}. Indeed, by day 3, populations expressing low levels of all fibroblast markers except CD47 enriched for reprogramming populations compared with highly expressing populations (Fig. 2a). Surprisingly though, at these early time points, cells with high levels of the ESC markers SSEA1, CD54, CD326 or CD71 did not show significantly increased reprogramming (Fig. 2a–c). Day 9 fractions expressing high levels of CD326 or SSEA1 began to show greater but insignificant enrichment for reprogramming populations. We confirmed previous reports that SSEA1-sorted cells produce more iPS cell colonies in mass culture^{1,8,10}, emphasizing the critical importance of the 96-well assay (Extended Data Fig. 4b, c). Unlike what was previously assumed, our findings demonstrate that acquisition of markers that define the pluripotent state is a late event and early expression of ESC markers has little predictive value for successful reprogramming. Additionally these data support the idea that the mesenchymal-to-epithelial transition, as judged by the epithelial marker CD326 (EpCAM), is a late event^{2,15–18,20}.

Our surface marker screen identified several specific markers for stable, partially reprogrammed cells^{4,9,18,21} (Fig. 1a). Although thought to be ‘stuck’ during reprogramming^{4,18,21}, we hypothesized a productive intermediate might arise between fibroblasts and iPS cells that share a subset of these markers. Indeed, day 6 fractions expressing high levels of the PRC markers CD73, CD49d and CD200 significantly enriched for reprogramming populations (Fig. 2b). Also, on day 9 the CD73^{high}

¹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, California 94305, USA. ²Department of Genetics, Stanford University, Stanford, California 94305, USA.

³Department of Pathology, Stanford University, Stanford, California 94305, USA. ⁴Baxter Laboratory in Stem Cell Biology, Department of Microbiology and Immunology, Stanford University, Stanford, California 94305, USA. ⁵Department of Microbiology and Immunology, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

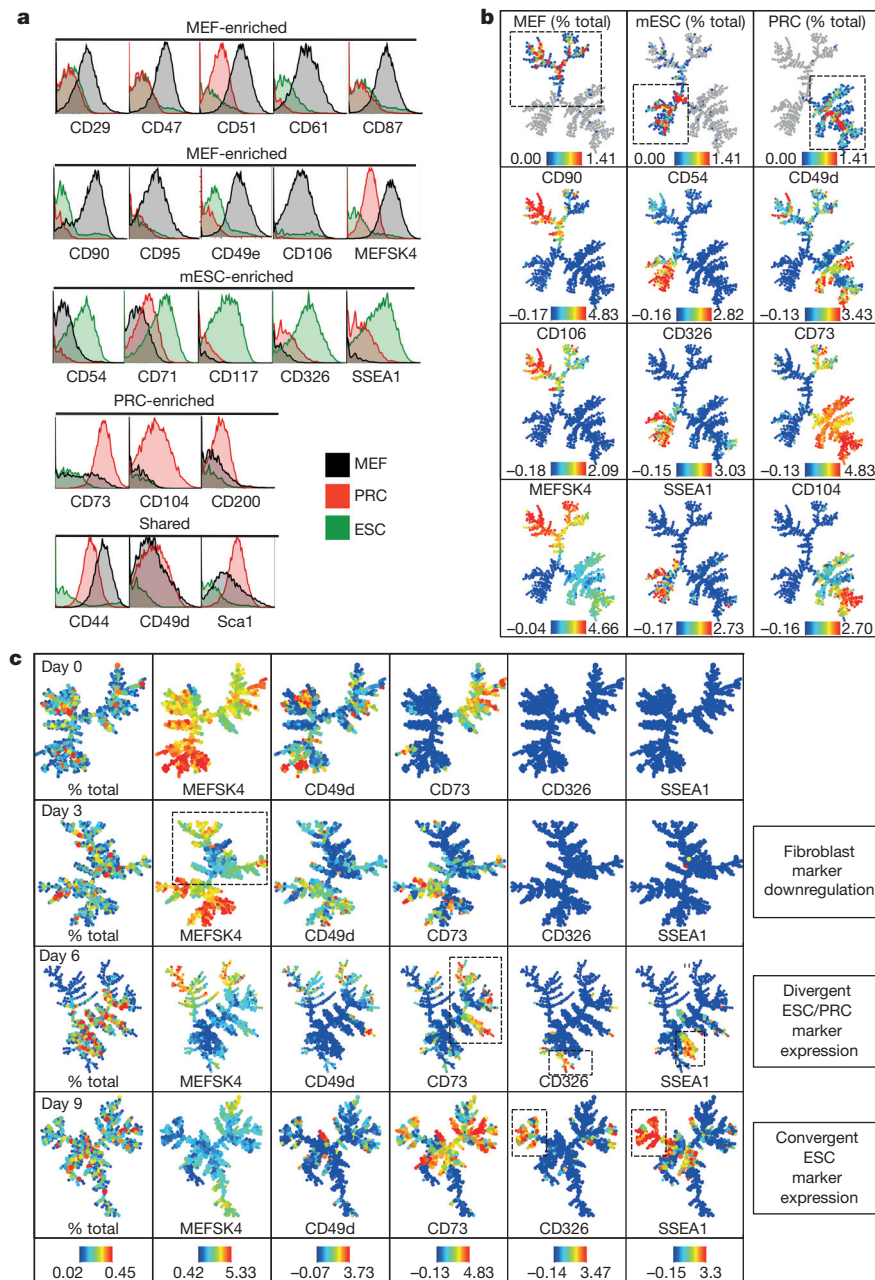


Figure 1 | Reprogramming surface marker profiling by mass cytometry.

a, Histogram overlays show mass cytometry signal intensity for the three cell populations analysed: mouse embryonic fibroblasts (MEF, black), partially reprogrammed cells (PRC, red) and embryonic stem cells (ESC, green).

b, SPADE analysis of combined MEF, PRC and ESC data sets. Colour bars

represent absolute percentages (top row) and ArcSinh-transformed counts for each marker. **c**, SPADE analysis of infected reprogramming MEF populations at days 0, 3, 6 and 9. For each marker, the same colour scale was applied to every sample, allowing direct comparison between time points. Colour bars represent absolute percentages (left) and ArcSinh-transformed counts.

and CD49d^{high} populations contained higher reprogramming activity (Fig. 2c). In agreement with these results, the day 6 SPADE analysis showed CD73^{high}, CD49d^{high} and CD200^{high} branches largely clustering independently from branches enriched for ESC and MEF markers (Fig. 1c). These results demonstrate that distinct intermediate populations arise after fibroblast program repression but before ESC marker acquisition.

We next focused on the markers CD73 and CD49d. When corrected for plating efficiency, both CD73^{high} and CD49d^{high} populations showed remarkably high reprogramming efficiencies of $9.5\% \pm 3.5$ and $12.5\% \pm 5.7$, respectively (Fig. 2d–f). Similar enrichment of a reprogramming-prone population was observed in reprogramming tail tip fibroblasts and glial-restricted neural precursor cells, suggesting CD73 and CD49d may be universal markers of intermediate reprogramming stages (Fig. 3a–d).

Finally, we explored their potential functional implications during reprogramming, and observed that adenosine, the enzymatic product of CD73, has a negative effect throughout and that CD49d activity is necessary during late reprogramming (Extended Data Fig. 4e).

We then used our day 6 SPADE analysis to identify heterogeneously expressed markers that could subdivide the CD73^{high} reprogramming-prone population and conducted single-cell efficiency assays (Fig. 3e, f). Within the CD73-positive population, the CD44^{high}, CD71^{low} and CD326^{high} fractions failed to reprogram, while CD49d^{high} and CD326^{low} fractions enriched for a reprogramming population. Thus, a CD73^{high} CD49d^{high} CD326^{low} CD44^{low} signature best describes the population undergoing productive reprogramming on day 6. Overlaying this signature onto the day 6 SPADE tree allowed determination of the exact cellular clusters most similar to this reprogramming-prone signature

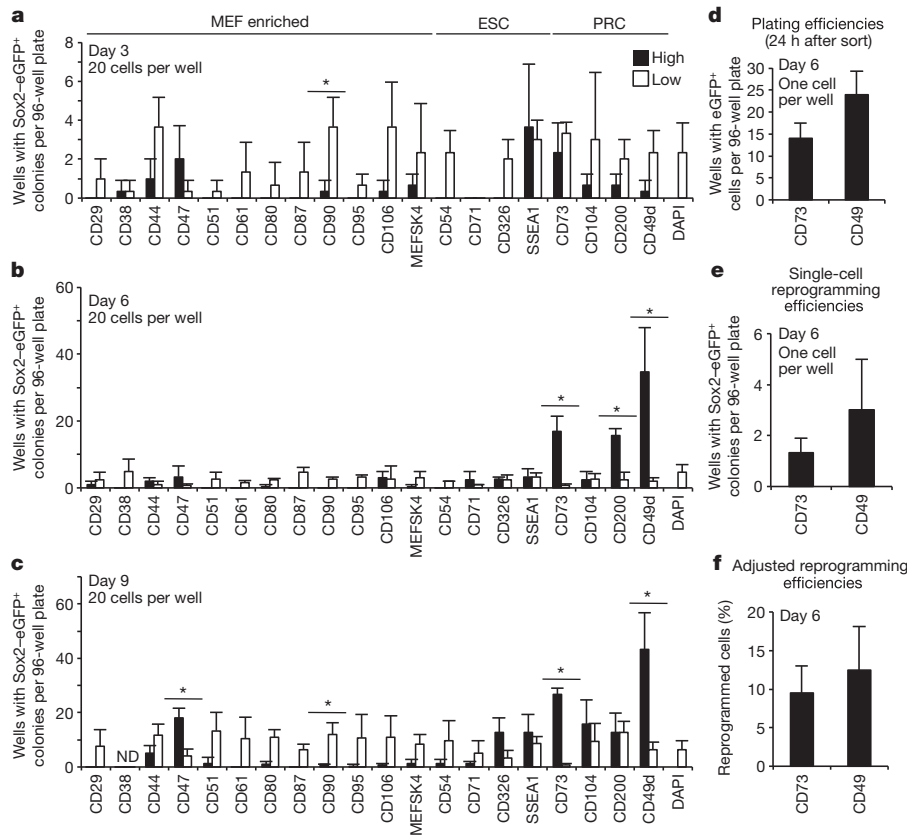


Figure 2 | A surface marker screen identifies an early CD73^{high} CD49d^{high} reprogramming intermediate. a–c, Ninety-six-well reprogramming assays on days 3, 6 and 9. Twenty cells per well sorted at days 3, 6 and 9. Sox2–eGFP⁺ colonies were assayed on day 24. Asterisks indicate two-sided *t*-test *P* < 0.05. d, Plating efficiencies for enhanced green fluorescent protein

(eGFP)-expressing MEFs sorted on day 6 for CD73^{high} or CD49d^{high} (assayed 24 h after sort). e, Single-cell reprogramming efficiencies for day 6 CD73^{high} or CD49d^{high} fractions. f, Reprogramming efficiencies (e) adjusted for plating efficiencies (d). Error bars, s.d.; *n* = 3 independent experiments for all assays.

(Fig. 3g). As expected, MEF markers were low in these poised populations. This suggests that this intermediate arises after loss of mesenchymal markers, but before completion of mesenchymal-to-epithelial transition, as indicated by reprogramming enrichment in the CD326^{low} fraction.

To identify subsequent reprogramming stages, we conducted continuation analysis where cells were sorted on day 6 and characterized by mass cytometry on day 16 (Fig. 3h and Extended Data Fig. 4–6). By day 10, reprogramming-prone populations formed distinct colonies with ESC-like morphology, while CD73^{low} cells were highly proliferative but failed to develop into mature colonies (Extended Data Fig. 5b). Continuation analysis on day 16 revealed that while reprogramming-prone and non-prone populations contained CD326-expressing cells, broad overlap between CD326^{high} and SSEA1^{high} clusters was only in mature reprogramming-prone populations (Fig. 3h and Extended Data Fig. 6). These clusters did not overlap with the ESC marker CD54, and were heterogeneous for CD73 and CD49d. We conclude a distinct CD326^{high}, SSEA1^{high}, CD54^{low} intermediate arises after the CD73^{high}/CD49d^{high} intermediate and before pluripotency acquisition.

We then used the intermediates stages to gain molecular insights into transcriptional regulation of early reprogramming. Gene expression analysis intriguingly showed these intermediates precede activation of the majority of transcription factors thought to be predictive markers for pluripotency induction (Fig. 4a and Extended Data Fig. 7)^{2,7}. The observation that these intermediates arise before key pluripotency regulators suggests a separate combination of early transcription factors must be induced to generate early intermediates and poise them for pluripotency acquisition. We found the transcription factors *Nr0b1* and *Etv5* preferentially expressed in reprogramming-prone populations and

highly expressed in ESCs, suggesting a functional role in poising early reprogramming (Fig. 4a and Extended Data Fig. 7b–d).

To assess whether these genes were necessary to induce the early intermediate populations, we generated three short hairpins against each gene (Extended Data Fig. 7e, f). We then assessed the ability of reprogramming MEFs infected with these short hairpins to induce CD73^{high}/CD49d^{high} intermediates 9 days after reprogramming induction (Fig. 4b and Extended Data Fig. 7g). Surprisingly, while MEFs infected with a control short hairpin were able to induce the CD73^{high}/CD49d^{high} intermediate (6.70 ± 2.27%), reprogramming MEFs infected with short hairpins targeting *Etv5* or *Nr0b1* were significantly impaired (Fig. 4b). This phenotype could be rescued by complementary DNA (cDNA) overexpression in combination with a hairpin targeting the untranslated region for the gene of interest (Fig. 4c). Further, when Nanog⁺ colonies or Sox2–eGFP⁺ colonies were assessed 24 days after reprogramming induction, a dramatic decrease in reprogramming efficiencies was observed (Fig. 4d, e). In contrast, knockdown of either gene in ESCs did not affect survival or proliferation (Extended Data Fig. 7h). While this paper was under review, an independent report confirmed *Nr0b1* as necessary for reprogramming²⁴. These data indicate that *Etv5* and *Nr0b1* are required to generate the CD73^{high}/CD49d^{high} poised intermediate necessary to induce the canonical pluripotency program and definitive iPS cell formation.

We then wondered whether a similar intermediate population arises in high-efficiency reprogramming systems^{12,13}. Published expression analysis of two high-efficiency systems showed transient CD73 upregulation, suggesting the presence of a similar intermediate (Extended Data Fig. 8a, b). We then characterized one of these systems, the *Mbd3*^{fl/fl}-secondary MEF system, in greater detail¹². After confirming reported

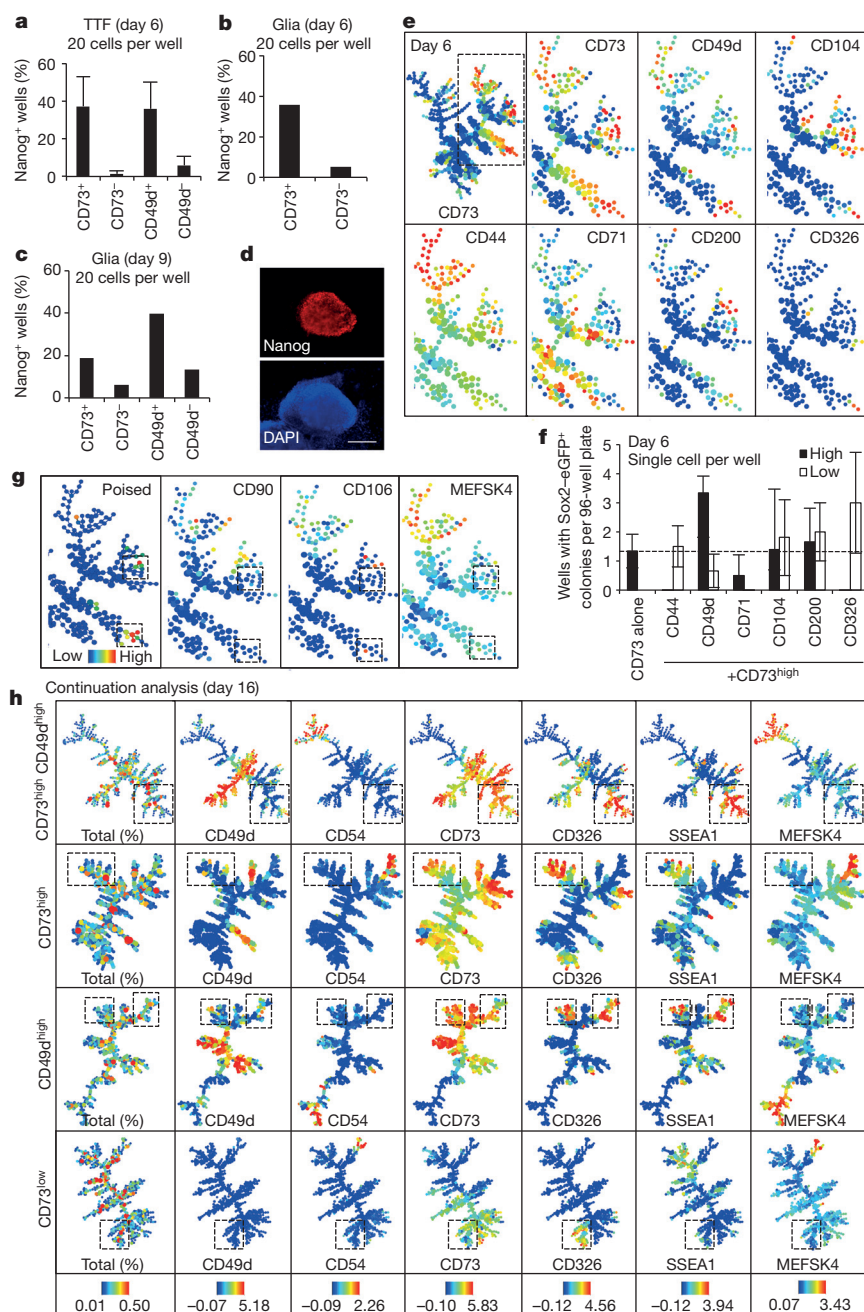


Figure 3 | Characterization of CD73^{high} and CD49d^{high} intermediates.

a–c, Percentage of wells with Nanog-expressing cells on day 24 for reprogramming tail tip fibroblasts (TTF) (**a**) ($n = 3$) and glia (**b**, **c**) ($n = 1$, independent primary cells and infections). **d**, Representative Nanog immunostaining. Scale bar, 200 μ m. **e**, Heterogeneously expressed markers in

day 6 CD73^{high} population. **f**, Single-cell 96-well assays for day 6 CD73^{high} fraction with additional surface markers ($n = 3$). **g**, Refined poised signature. Clusters are low for mesenchymal markers. **h**, Continuation analysis shows SSEA1^{high} CD326^{high} branch unique to poised populations (boxed). All experiments represent independent biological replicates. Error bars, s.d.

reprogramming efficiencies, we analysed this system by mass cytometry (Extended Data Figs 8–10). By day 3, fibroblast marker repression was evident, and CD73 was upregulated within this population (Extended Data Figs 8e and 9b). Within the CD73^{high}/MEFSK4^{low} population, CD49d (*Itga4*) upregulation was not apparent, but we noticed the emergence of a separate integrin, CD104 (*Itgb4*). By day 4, the major CD73^{high} branch clearly overlapped with the CD104^{high} branch and persisted into day 5. SSEA1^{high} and CD326^{high} expression was present on day 4, but clear co-expression was not seen until day 5. By day 9, CD73 and CD104 expression was dramatically reduced while CD326 and SSEA1 expression remained high. These data demonstrate a transient CD73^{high}/CD104^{high} population arises after donor cell program repression and

before ESC marker acquisition, even in a highly efficient reprogramming system. Similar to CD49d and CD73, CD104 is not highly expressed in ESCs (Fig. 1a). And similar to viral reprogramming, adenosine treatment abolished reprogramming in the Mbd3 reprogramming system, albeit only at late stages, whereas compounds affecting CD49d function had little effect (Extended Data Fig. 4f).

The stage-specific framework provided in this study bridges the previously unexplored gap between donor program silencing and pluripotent marker acquisition (Fig. 4f). We demonstrate a transient, 'poised' intermediate present across multiple reprogramming systems, suggesting a general property of iPSC reprogramming. We note that, similar to SSEA1, TRA-1-60 enriches for reprogramming-prone intermediates at

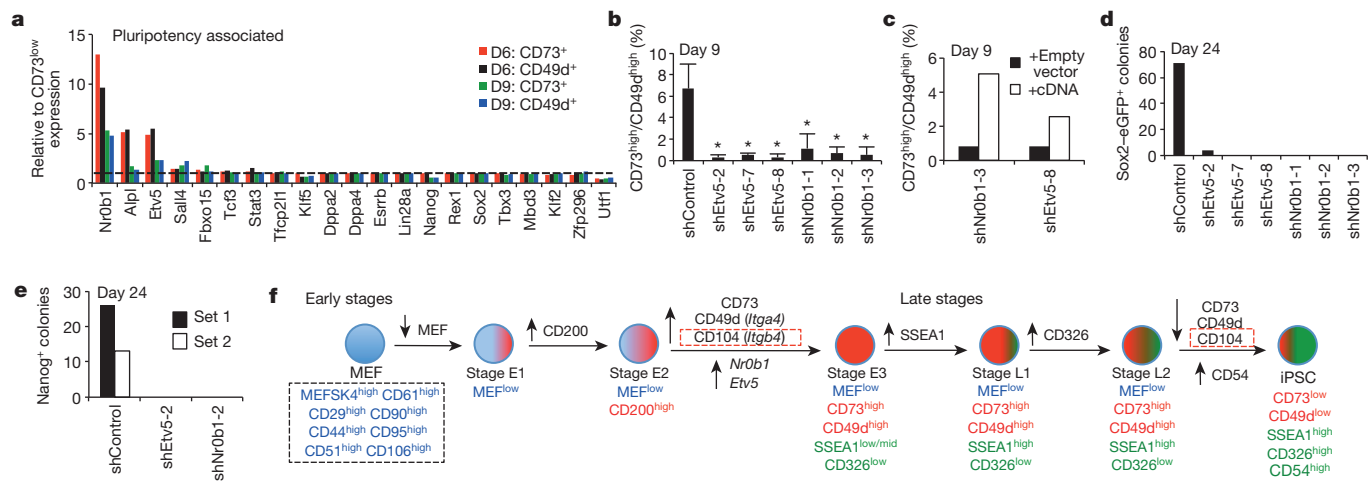


Figure 4 | Reprogramming regulators identified with CD73^{high}/CD49d^{high} intermediates. **a**, Day 6 and 9 reprogramming-prone and non-prone pluripotency-associated gene differential expression. Dotted line represents value of 1 (no difference). **b**, **c**, Day 9 CD73^{high}/CD49d^{high} quantification for knockdown (**b**, $n = 3$ independent experiments) and rescue (**c**) experiments. Gating shown in Extended Data Fig. 7g. Asterisks, two-sided t -test

later time points during human reprogramming^{11,25}, and we speculate a similar transient intermediate arises in the human system.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 March 2014; accepted 2 February 2015.

Published online 1 April; corrected online 20 May 2015 (see full-text HTML version for details).

1. Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151–159 (2008).
2. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–1222 (2012).
3. Golipour, A. *et al.* A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* **11**, 769–782 (2012).
4. Chen, J. *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genet.* **45**, 34–42 (2013).
5. Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).
6. Ho, R., Papp, B., Hoffman, J. A., Merrill, B. J. & Plath, K. Stage-specific regulation of reprogramming to induced pluripotent stem cells by Wnt signaling and T cell factor proteins. *Cell Rep.* **3**, 2113–2126 (2013).
7. O'Malley, J. *et al.* High-resolution analysis with novel cell-surface markers identifies routes to iPSCs. *Nature* **499**, 88–91 (2013).
8. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPSCs. *Cell* **151**, 1617–1632 (2012).
9. Sridharan, R. *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136**, 364–377 (2009).
10. Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPSC cell reprogramming in mouse. *Cell Stem Cell* **2**, 230–240 (2008).
11. Tanabe, K., Nakamura, M., Narita, M., Takahashi, K. & Yamanaka, S. Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc. Natl Acad. Sci. USA* **110**, 12172–12179 (2013).
12. Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).
13. Di Stefano, B. *et al.* C/EBP α poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature* **506**, 235–239 (2014).
14. Hou, P. *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651–654 (2013).
15. Zunder, E. R., Lujan, E., Goltsev, Y., Wernig, M. & Nolan, G. P. A continuous molecular roadmap to iPSC reprogramming through progression analysis of single cell mass cytometry. *Cell Stem Cell* **16**, 323–337 (2015).

$P < 0.05$. **d**, **e**, Day 24 Sox2-eGFP⁺ (**d**, $n = 1$) and Nanog⁺ colonies (**e**, $n = 2$ independent experiments) from *Rosa-rtTA*[±], *Sox2-eGFP*[±] and *Rosa-rtTA*[±] MEFs, respectively. **f**, Early and late reprogramming model. Dotted red boxes distinguish CD104 observed in the *Mbd3*^{fl/-} system. Error bars, s.d.

16. Ichida, J. K. *et al.* A small-molecule inhibitor of Tgf- β signaling replaces Sox2 in reprogramming by inducing *Nanog*. *Cell Stem Cell* **5**, 491–503 (2009).
17. Li, R. *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* **7**, 51–63 (2010).
18. Maherali, N. & Hochedlinger, K. Tgf β signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Curr. Biol.* **19**, 1718–1723 (2009).
19. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
20. Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
21. Meissner, A., Wernig, M. & Jaenisch, R. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nature Biotechnol.* **25**, 1177–1181 (2007).
22. Bandura, D. R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal. Chem.* **81**, 6813–6822 (2009).
23. Qiu, P. *et al.* Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nature Biotechnol.* **29**, 886–891 (2011).
24. Zhang, J. *et al.* Dax1 and Nanog act in parallel to stabilize mouse embryonic stem cells and induced pluripotency. *Nature Commun.* **5**, 5042 (2014).
25. Takahashi, K. *et al.* Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nature Commun.* **5**, 3678 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank P. Lovelace, R. Finck, K. M. Loh, K. Tanabe and S. Marro for advice. We thank J. Hanna for his the secondary *Mbd3*^{fl/-} MEFs. We also thank S. Knöbel for the mEF-SK4 antibody. This work was supported by the California Institute of Regenerative Medicine grant RB2-01592 (to G.P.N.), the Institute for Stem Cell Biology and Regenerative Medicine at Stanford, and a New York Stem Cell Foundation-Robertson Investigator Award. E.L. was supported by the California Institute for Regenerative Medicine Predoctoral Fellowship TG2-01159 and National Science Foundation Graduate Research Fellowship DGE-114747. E.R.Z. was supported by National Institutes of Health National Research Service Award F32 GM093508-01. M.W. is a New York Stem Cell Foundation-Robertson Investigator and a Tashia and John Morgridge Faculty Scholar at the Child Health Research Institute at Stanford.

Author Contributions E.L., E.R.Z., G.P.N. and M.W. designed research. E.L. conducted reprogramming and sorting experiments. E.R.Z. conducted mass cytometry analysis and data processing. Y.H.N. and I.N.G. assisted with sample processing. E.L., E.R.Z., G.P.N. and M.W. analysed data. E.L., E.R.Z., G.P.N. and M.W. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.W. (wernig@stanford.edu).

METHODS

No statistical methods were used to predetermine sample size.

Cell culture. Embryonic fibroblasts were isolated from embryonic day (E)13.5 embryos derived from B6;129S-*Sox2*^{eGFP/+} mated with B6.Cg-Gt(*ROSA*)26Sor^{tm1(rtTA⁺M2)} (allele referred to as *Rosa-rtTA*) as previously described^{26,27}. Tail tip fibroblasts were derived from 1-week-old B6 mice. This was done by chilling the animals on ice for 5 min, cutting the tail, and mincing in a 6 cm dish in 1 ml 0.25% trypsin. Another 1 ml of 0.25% trypsin was added after mincing and this was incubated at 37 °C for 10 min. The mixture was then resuspended in 15 ml MEF medium (described below) and plated on a 0.2% gelatinized 15 cm plate. Glial cultures were prepared from CD1 mice as previously described²⁶. MEFs, tail tip fibroblasts and glial cells were cultured in MEF medium, which consisted of 10% cosmic calf serum (Thermo Scientific) in DMEM (Invitrogen) supplemented with non-essential amino acids (Invitrogen), penicillin–streptomycin (Invitrogen), sodium pyruvate (Invitrogen) and 2-mercaptoethanol (Invitrogen). Primary *Mbd3*^{fl/+}, *Oct4-GFP* secondary MEFs¹² were a gift from J. Hanna and grown in MEF medium. We verified modification of the *Mbd3* locus by western blot analysis (described below under ‘Western blotting’; Extended Data Fig. 8f). Our partially reprogrammed line was previously characterized^{15,21}. These were grown in MEF medium supplemented with 15% cosmic calf serum. *Oct4-neoR* knock-in mouse ESCs^{19,21} were grown in mESC medium, which consisted of 12% knockout replacement serum (Invitrogen), 3% cosmic calf serum and supplemented with non-essential amino acids, penicillin–streptomycin, sodium pyruvate, 2-mercaptoethanol and leukaemia inhibitory factor (LIF). MEFs, PRCs and mESCs were routinely tested for mycoplasma contamination.

Surface marker screening. We screened mouse surface markers using the ‘Mouse cell surface marker screening panel’ Lyoplate (BD Biosciences, material number 562208) according to the manufacturer’s instructions. Briefly, 150 × 10⁶ cells were used for our partially reprogrammed cells (passage 15) and *Oct4-neoR* mES cells. Cells were plated on 0.2% gelatin-coated plates and treated with neomycin 3 days before screening; 175 × 10⁶ cells were used for passage 4 *Sox2-eGFP* MEFs. Cells were washed with PBS-EDTA, dissociated in 10× TrypLE (Invitrogen) for 5 min, washed once in PBS before staining for 30 min in primary antibody in staining medium (PBS-EDTA supplemented with 0.5% BSA) according to manufacturer’s instructions. After the primary antibody stain, cells were washed two times in PBS, stained for 30 min. on ice with biotinylated secondary antibodies in staining medium, washed two times in PBS, and then stained for 30 min. on ice with Alexafluor 647-conjugated streptavidin in staining medium. After the tertiary Alexafluor 647–streptavidin stain, the cells were washed twice in PBS and analysed on a BD LSR II Flow Cytometer with 96-well HTS module in staining medium. We used a high concentration of TrypLE to verify that all identified markers would not be cleaved by our dissociation reagent. We further validated a subset of these identified makers with our control populations dissociated in 1× TrypLE and stained with fluorophore-conjugated antibodies as described below.

Reprogramming. Reprogramming assays were conducted with passage 4 *Rosa-rtTA*⁺, *Sox2*^{eGFP/+} or *Rosa-rtTA*⁺ (derived from the same litter: see mating scheme above) mouse embryonic fibroblasts, passage 2 B6 tail tip fibroblasts or passage 3 B6 glia as indicated. FUW-tetO lentiviral vectors and lentiviral packaging were used as previously described²⁶. Passage 3 MEFs and passage 2 glial cells were split onto 10 cm 0.2% gelatin-coated plates at 250,000 cells per plate 1 day before infection. Passage 1 tail tip fibroblasts were split onto 10 cm 0.2% gelatin-coated plates at 100,000 cells per plate 1 day before infection. Cells were infected in MEF medium supplemented with polybrene (8 µg ml⁻¹; Sigma). One day after infection, MEF medium supplemented with doxycycline was added; this was considered to be day 0. Media were replaced every 2 days. On day 16, medium was replaced with mES cell medium without doxycycline. Owing to the complication of getting a large number of glial cells, we performed the reprogramming efficiency assays once per time point (days 6 and 9) but both were done with independently derived primary cells from different animals, independently derived virus and sorts.

The following compounds were also added to reprogramming medium on the days indicated in the text: BIO 1211 (Tocris, 4 nM), adenosine (Sigma, 20 µM or 2 µM as indicated), fibronectin (Sigma, 4 µg ml⁻¹), AMP-CP (adenosine 5′-(α,β-methylene)diphosphate) (Sigma, 20 µM). These were chosen as CD73 (5′-nucleotidase, ecto) converts AMP to adenosine, CD49d (*Itga4*) is an integrin whose substrates are VCAM1 and fibronectin, α,β-methyleneadenosine 5′-diphosphate (AMP-CP) is a CD73 inhibitor and BIO1211 is a CD49d inhibitor.

Hairpins for *Nr0b1* and *Etv5* were designed with the pSico oligomaker (Supplementary Table 3) and cloned into the lentiviral pSico-puro vector²⁸ (<http://web.mit.edu/jacks-lab/protocols/pSico.html>). We assessed knockdown efficiencies in partially reprogrammed cells as they express these genes and grow well in MEF medium supplemented with 15% serum alone. To assess knockdown efficiencies, 50,000 PRCs were plated onto a six-well gelatin-coated well 1 day before transduction. Cells were transduced with the hairpin of interest in MEF medium supplemented

with 15% serum and polybrene (8 µg ml⁻¹). The medium was exchanged the following day with MEF medium supplemented with 15% serum and puromycin (2 µg ml⁻¹). Cells were cultured for a further 3 days before RNA extraction (for a total of 4 days after infection). RNA was prepared with an RNeasy purification kit as described below.

To assess the effects of knockdown of these genes on the day 9 CD73^{high}/CD49d^{high} intermediate, 100,000 P4 *Rosa-rtTA* MEFs were plated on a 10 cm gelatin-coated plate. These were transduced the following day with the doxycycline-inducible FUW-tetO-*Klf4*, *Oct3/4*, *Sox2* and *c-Myc* vectors and indicated hairpins as described above. Medium was supplemented with doxycycline 1 day after infection (day 0). We note that because *Rosa-rtTA* MEFs contain a PGK-puromycin cassette, they were not selected with puromycin. On day 9, CD73-Alexa 488 (1:50) and CD49d-Alexa 750 (1:50) were analysed by fluorescence-activated cell sorting (FACS staining described below). To assess reprogramming efficiencies for these genes, 30,000 P4 *Rosa-rtTA* MEFs were plated on a 6 cm gelatin-coated plate or 15,000 P4 *Rosa-rtTA*, *Sox2-eGFP* MEFs were plated on a six-well gelatin-coated well and transduced the following day in the same manner. At day 16, the culture medium was switched to mES cell medium without doxycycline. Twenty-four days after reprogramming induction, Nanog⁺ colonies were assessed for *Rosa-rtTA* MEFs (immunofluorescence staining described below) or Sox2-eGFP⁺ colonies for *Rosa-rtTA*, *Sox2-eGFP* MEFs. In total, three independent reprogramming efficient experiments were conducted across the two different MEF lines. To validate further the specificity of the effects, we performed a ‘rescue’ experiment to demonstrate that upon re-expression of the cDNA (under knockdown conditions) the effect is eliminated. To this end, the cDNAs for *Nr0b1* and *Etv5* were cloned into the FUW-tetO vector. Fifteen thousand P4 *Rosa-rtTA* MEFs on six-well gelatin-coated plates were then transduced with the doxycycline-inducible FUW-tetO-*Klf4*, *Oct3/4*, *Sox2* and *c-Myc* vectors and indicated hairpins with the FUW-tetO-cDNA (*Etv5* or *Nr0b1*) or an empty vector. Medium was supplemented with doxycycline 1 day after infection (day 0). On day 9, CD73-Alexa 488 (1:50) and CD49d-Alexa 750 (1:50) were analysed by FACS (FACS staining described below). As the rescue experiment was supplementary to and consistent with the main knockdown experiments, we conducted only one rescue experiment for one hairpin.

To verify that mouse ESCs could survive and proliferate after knockdown of *Etv5* or *Nr0b1*, we infected 30,000 ESCs per well in gelatinized six-well plates in mESC medium and replaced the medium the following day with mESC medium supplemented with puromycin to select for successful transduction. These were then cultured for 3 days, dissociated with 0.25% trypsin and re-plated onto gelatinized six-well plates. These were then cultured for 3 days, fixed and stained for *Oct4* (described below).

To reprogram *Mbd3*^{fl/+}, *Oct4-GFP* secondary MEFs (the *Oct4-GFP* allele in these cells is not a targeted, but well-characterized, transgenic reporter¹²), we first assayed reprogramming conditions with 2000 *Mbd3*^{fl/+} secondary MEFs on 2 × 10⁶ mitomycin- (Sigma) treated B6 feeders, 10⁵ *Mbd3*^{fl/+} secondary MEFs without feeders and 2 × 10⁵ *Mbd3*^{fl/+} secondary MEFs without feeders on 10 cm dishes coated with 0.2% gelatin. We found optimal reprogramming efficiencies with 2,000 *Mbd3*^{fl/+} secondary MEFs and 2 × 10⁶ feeders as previously reported¹². All reprogramming assays were done in non-hypoxic conditions. To reprogram cells, cells were cultured in media 1 (recombinant human LIF (10 ng ml⁻¹, Peprotech), doxycycline (1 µg ml⁻¹) and ascorbic acid (10 µg ml⁻¹, Sigma) in 15% cosmic calf serum (Thermo Scientific) in DMEM (Invitrogen) supplemented with non-essential amino acids (Invitrogen), penicillin–streptomycin (Invitrogen), sodium pyruvate (Invitrogen) and 2-mercaptoethanol (Invitrogen)) for 3 days and then media 2 (recombinant human LIF (10 ng ml⁻¹), doxycycline (1 µg ml⁻¹), PD0325901 (1 µM, Cell Signaling) and CHIR99021 (3 µM, Cayman) in 15% knockout replacement serum (Invitrogen) in DMEM (Invitrogen) supplemented with non-essential amino acids (Invitrogen), penicillin–streptomycin (Invitrogen), sodium pyruvate (Invitrogen) and 2-mercaptoethanol (Invitrogen)), until the completion of the experiment. Media were exchanged every 2 days.

Mass cytometry analysis. On designated days the reprogramming cultures were treated with 1× TrypLE (Invitrogen) for 5 min at 37 °C, dissociated into single-cell suspension by trituration, and then washed twice with PBS. The cell samples were then incubated with metal-conjugated antibodies (Supplementary Table 1) in PBS containing 5% FBS (Omega Scientific) for 30 min on ice, washed once with PBS containing 5% FBS, treated with 25 µM cisplatin for 1 min on ice for live–dead cell discrimination²⁹, washed once with PBS containing 5% FBS and then fixed with 1.6% paraformaldehyde at 20 °C (room temperature) for 10 min. Formaldehyde-fixed cell samples were then permeabilized with methanol on ice for 15 min, washed once with PBS containing 0.5% BSA, and then incubated at room temperature for 15 min with an iridium-containing DNA intercalator (DVS Sciences/Fluidigm) in PBS containing 1.6% paraformaldehyde. After intercalation/fixation, the cell samples were washed once with PBS containing 0.5% BSA and twice with water before measurement on a CyTOF mass cytometer (DVS Sciences/Fluidigm).

Normalization for detector sensitivity was performed as previously described³⁰, using polystyrene normalization beads containing lanthanum-139, praseodymium-141, terbium-159, thulium-169 and lutetium-175. Number of cells for each mass cytometry experiment are shown in Supplementary Table 2.

SPADE analysis. Density-dependent downsampling, hierarchical clustering, cluster upsampling and extraction of parameter medians was performed by the SPADE package (www.cytospade.org) as described in the main text and previously^{23,31}. All assayed surface markers were used in the clustering step unless otherwise indicated, and the parameters for downsampling percentile and target number of clusters were set to 5% and 500, respectively.

The refined poised signature shown in Fig. 3g was determined by calculating the similarity of each SPADE cluster to the hand-gated CD73^{high} CD49d^{high} CD44^{low} CD326^{low} population. Similarity was calculated by the Manhattan distance metric using all measured surface markers, and is indicated by the coloured scale bar (low distance equals high similarity).

Immunofluorescence. Plates were fixed in 4% PFA for 10 min, washed three times with PBS, blocked and permeabilized in PBS supplemented with 5% CCS and 0.1% Triton-X 100 (Sigma) (blocking solution) for 10 min. Ninety-six-well plates were then incubated with mouse anti-Nanog (1:500, BD) or mouse anti-Oct4 (1:200, Santa Cruz Biotechnology) in blocking solution for 30 min, washed three times with PBS, incubated with donkey anti-mouse Alexa-555 (1:1,000, Invitrogen) or anti-mouse Alexa-488 (1:1,000, Invitrogen) in blocking solution for 30 min, washed three times with PBS and stained with 4',6-diamidino-2-phenylindole (DAPI) for 3 min. Cells were then washed with PBS and visualized.

FACS and efficiency assays. Cells were washed with PBS-EDTA, dissociated in 1× TrypLE for 5 min, washed with PBS and incubated on ice with a fluorophore-conjugated antibody and DAPI for 30 min in PBS-EDTA supplemented with 0.5% BSA. The sources and detailed descriptions of all antibodies used are listed in Supplementary Table 1. For 96-well assays, single or 20 DAPI⁺ cells per well were double sorted on the indicated day into gelatinized 96-well plates supplemented with 400,000 feeders per plate in MEF medium supplemented with doxycycline. For the primary sort, cells were sorted into PBS supplemented with 0.5% BSA; for the secondary sort, cells were sorted directly into 96-well plates. Efficiency assays were conducted 24 days after transgene induction and determined by the number of wells with Sox2-eGFP⁺ colonies. For mass culture reprogramming, 10,000 SSEA1^{high} or SSEA1^{low} cells were double sorted 6 days after transgene induction onto 3 cm gelatinized plates supplemented with feeders, and Sox2-eGFP⁺ colonies were assayed 24 days after transgene induction. We used the same SSEA1 clone and vendor as previously used for determining reprogramming efficiencies^{8,10}. Tail tip fibroblast and glial reprogramming efficiencies were determined by double sorting on the indicated days into 96-wells (as described above) and assaying for Nanog by immunofluorescence. Plating efficiencies were determined by infecting Sox2-eGFP MEFs with FUW-tetO-hygroB-T2A-eGFP, selecting for 5 days in hygromycin, and counting the number of wells with GFP⁺ cells 24 h after sorting. For all assays where CD73 was used for sorting, CD73-Alexa 647 was used (gating shown in Extended Data Fig. 4). For CD73 × CD49d analysis, CD73-Alexa 488 was used (Extended Data Fig. 7g).

RNA preparation and expression analysis. RNA of reprogramming populations for microarray analysis was prepared from *Rosa-rtTA*[±] day 6 and day 9 reprogramming cultures double sorted for CD73 or CD49d (as described above). RNA of con-

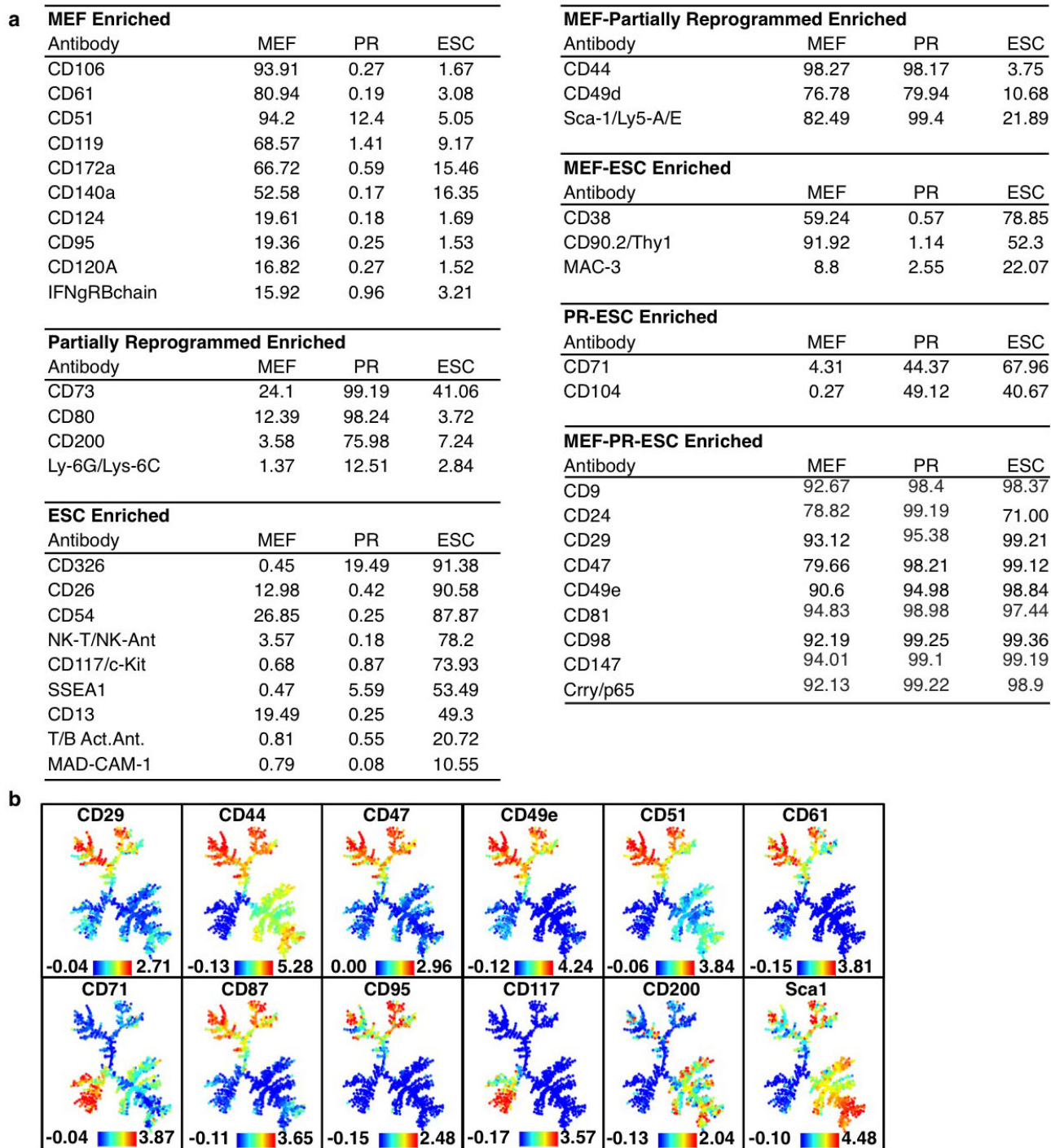
trol populations for microarray analysis was prepared from *Rosa-rtTA*[±] MEFs (passage 4), partially reprogrammed cells (passage 10) and V6.5 mouse ESCs (passage 11). RNA was prepared with RNeasy Mini Kit (Qiagen) and DNA was removed by on-column RNase-Free DNase treatment (Qiagen) according to the manufacturer's instructions. Mouse Gene 2.0 ST Arrays (Affymetrix) were prepared by the Stanford Protein and Nucleic Acid Facility. Data were normalized and gene names were assigned by Partek Genomic Suite. For all analysis, non-coding transcripts were removed. Preprocessing (floor = 100, ceiling = 20,000, min fold change = 2), *k*-means clustering (*k* = 5, seed value = 12345) and hierarchical clustering of *k*-means clusters (Pearson correlation, pairwise complete-linkage), and heat maps were generated by Gene Pattern (<http://www.broadinstitute.org/cancer/software/genepattern/>). Microarray data can be accessed with accession number GSE62957 from the National Center for Biotechnology Information database.

For Fig. 4a, genes were selected on the basis of pluripotency-associated genes characterized in previous studies^{2,7} or from differential expression of sorted populations and ESCs. *Oct4* is not shown as the probe failed to detect expression in mESCs.

For quantitative PCR analysis, cDNA was generated with a SuperScript First-Strand Synthesis System (Invitrogen). Data were generated with a 7900HT Real-Time PCR System (Applied Biosystems). Six-microlitre reactions were prepared with SYBR Green Real-Time PCR Master Mix (Life Technologies) under the following conditions: 50 °C for 2 min, 95 °C for 10 min, and 40 cycles of 15 s at 95 °C and 1 min at 60 °C. All expression was normalized to GAPDH before comparing with control hairpin expression levels. Supplementary Table 3 gives primer sequences.

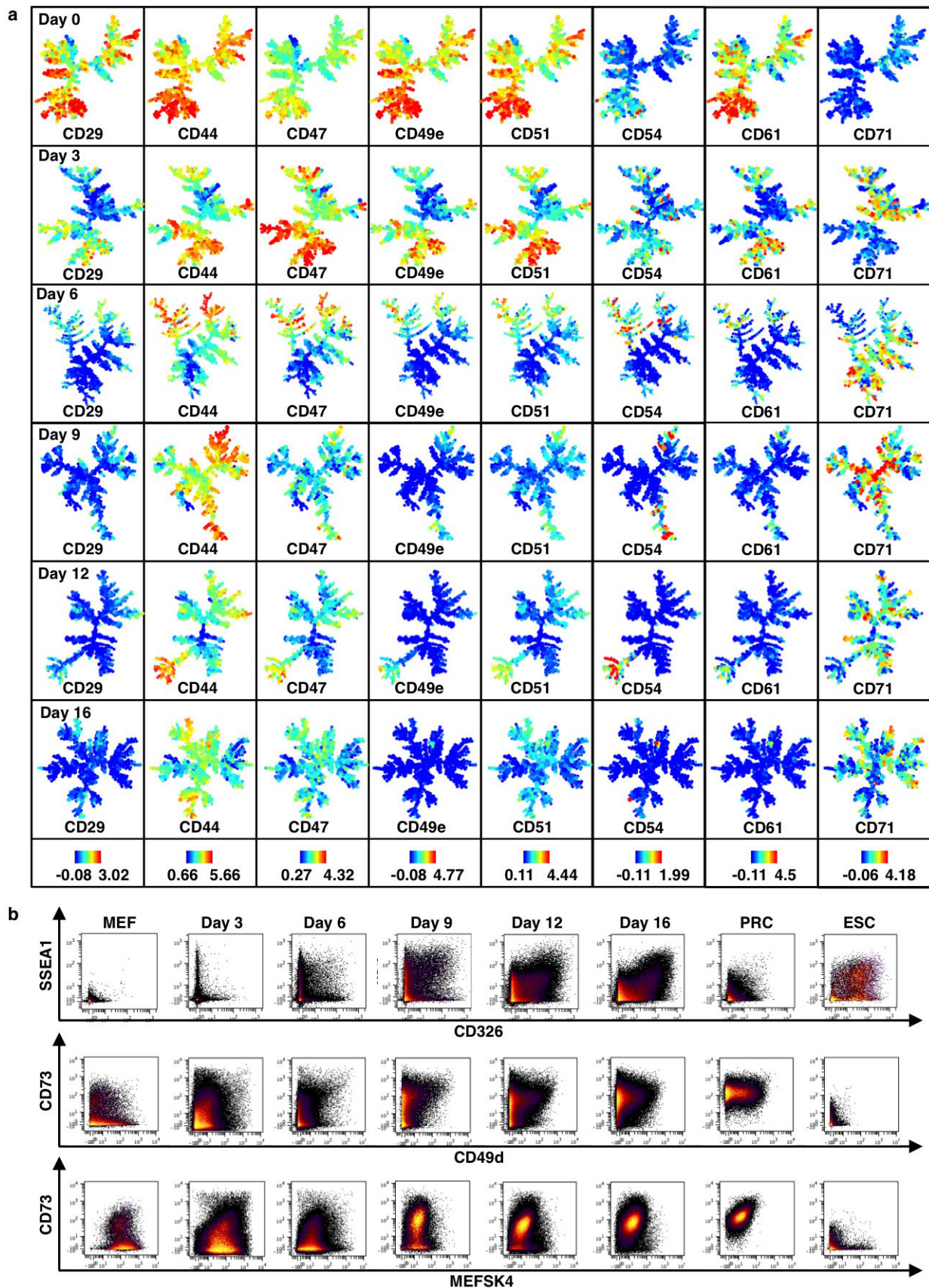
Western blotting. Passage 3 *Rosa-rtTA*[±] MEFs or secondary *Mbd3*^{fl/-}, *Rosa26-CreER* MEFs were grown in 10-cm tissue culture plates. Secondary *Mbd3*^{fl/-} MEFs were treated with 1 μM 4OH-tamoxifen for 24 h, then samples were cultured for a further 48 h and dissociated with 0.25% trypsin. The cell pellet was then lysed with cell lysis buffer (200 mM NaCl, 50 mM Tris pH8.0, 1% Triton X-100, 5% glycerol). Twenty micrograms of soluble protein was run on a 4–12% gradient Bis-Tris gel (Life Technologies) and blotted onto PVDF membrane. After blocking, membrane was incubated with primary antibody against Mbd3 (1:1,000, Bethyl Laboratories, A302-528A) for 1 h at room temperature, washed three times with PBS with 0.1% Tween-20 and incubated with secondary antibody anti-rabbit HRP (1:5,000, Jackson Immuno, A5441) for 1 h at room temperature.

26. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
27. Ellis, P. *et al.* SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* **26**, 148–165 (2004).
28. Robinson, D. A. *et al.* A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nature Genet.* **33**, 401–406 (2003).
29. Fienberg, H. G., Simonds, E. F., Fantl, W. J., Nolan, G. P. & Bodenmiller, B. A platinum-based covalent viability reagent for single-cell mass cytometry. *Cytometry* **81**, 467–475 (2012).
30. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry* **83**, 483–494 (2013).
31. Linderman, M. D. *et al.* CytoSPADE: high-performance analysis and visualization of high-dimensional cytometry data. *Bioinformatics* **28**, 2400–2401 (2012).



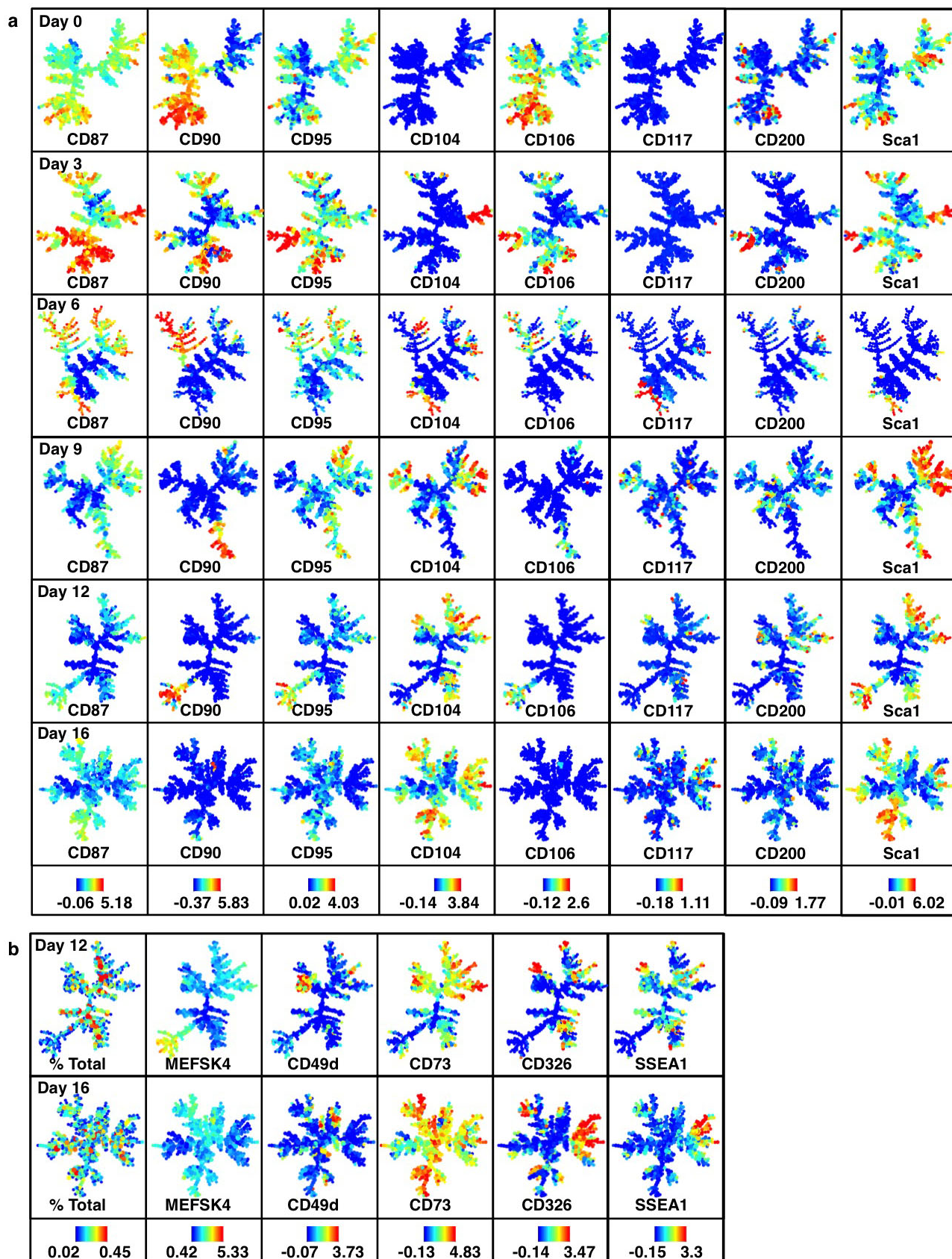
Extended Data Figure 1 | Results from surface marker screen. **a**, Shown are surface markers detected in MEFs, partially reprogrammed cells (PR) or ESCs analysed by flow cytometry. Numbers indicate the percentage of each population positive for the marker of interest, relative to isotype control

samples. Markers are grouped for enrichment in single populations or shared between multiple populations. **b**, SPADE analysis for MEFs, mESCs and PRCs for surface markers analysed by mass cytometry (continued from Fig. 1b). Colour bars (bottom) represent ArcSinh-transformed counts for each marker.



Extended Data Figure 2 | SPADE and biaxial analysis for MEF reprogramming. **a**, SPADE analysis of lentiviral-infected MEF reprogramming populations analysed by mass cytometry (continued from

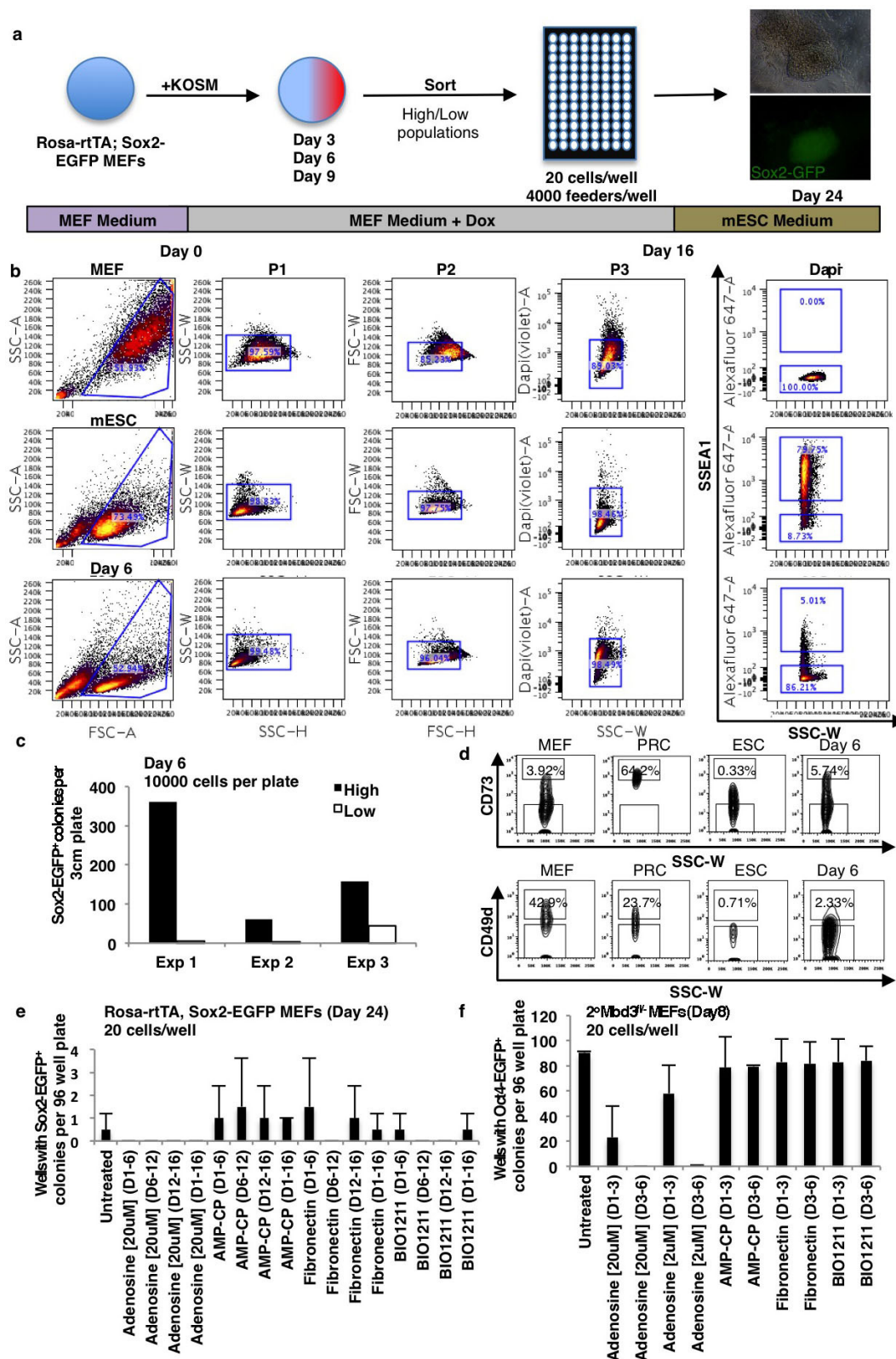
Fig. 1c). Colours bars (bottom) represent ArcSinh-transformed counts for each marker. **b**, Biaxial plots for selected markers in control populations and during MEF reprogramming.



Extended Data Figure 3 | SPADE analysis for MEF reprogramming.

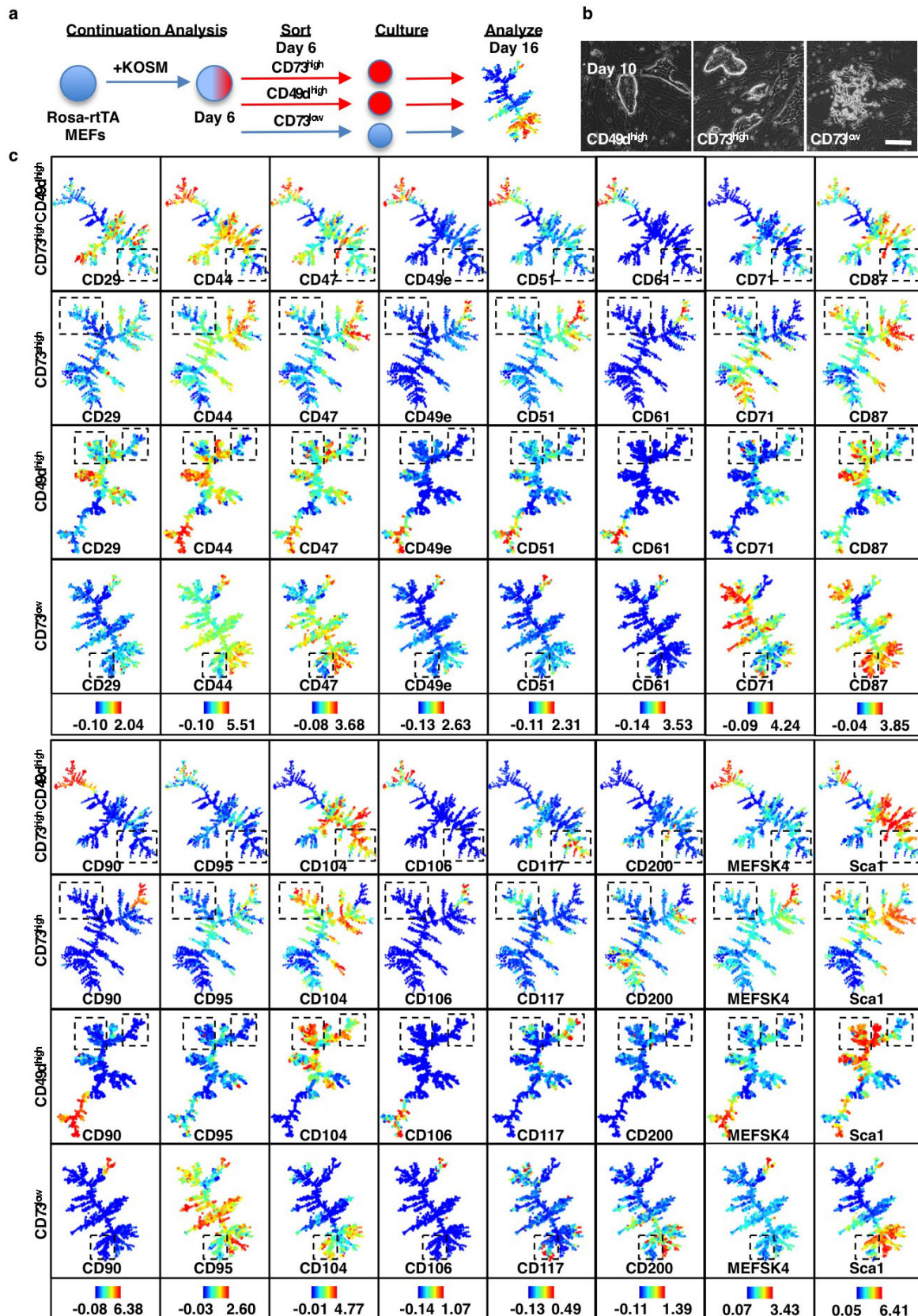
a, SPADE analysis of lentiviral-infected MEF reprogramming populations analysed by mass cytometry (continued from Fig. 1c). Colours bars

(bottom) represent ArcSinh-transformed counts for each marker. **b**, Day 12 and 16 time points for markers shown in Fig. 1c). Coloured bars for percentage total represent absolute percentages.



Extended Data Figure 4 | Details for sorting experiments and chemical treatment assay. **a**, Ninety-six-well reprogramming assay. Twenty cells per well sorted at days 3, 6 and 9. Sox2-eGFP⁺ colonies were assayed on day 24. **b**, Gating strategy for SSEA1 in controls and day 6 reprogramming population. High- and low-expressing populations were determined on the basis of MEF and ESC control levels. **c**, Ten thousand SSEA1^{high} (black bar) or SSEA1^{low} (white bar) were sorted onto 3 cm gelatinized plates with feeders. Sox2-eGFP⁺ colonies were counted on day 24 ($n = 3$ independent experiments). **d**, Gating strategy for CD73 and CD49d in controls and day 6

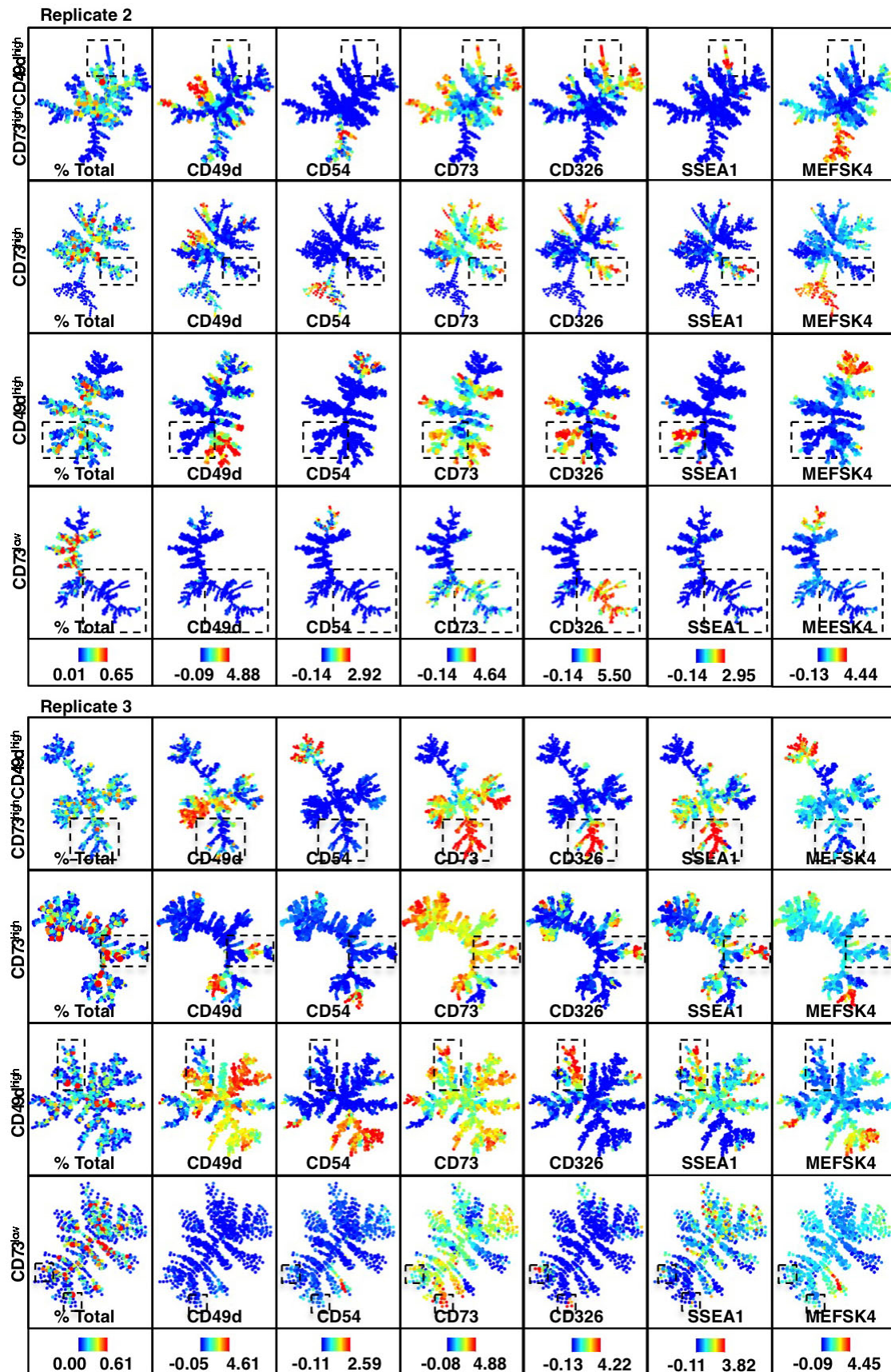
reprogramming population. High- and low-expressing populations were determined on the basis of MEF, PRC and ESC control levels. **e**, **f**, Treatment of reprogramming populations with compounds affecting CD73 and CD49d. Shown are 96-well reprogramming efficiencies for infected Rosa-rtTA Sox2-eGFP MEFs (**e**) or secondary *Mbd3*^{fl/-} MEFs (**f**). The y axis displays wells with Sox2-eGFP⁺ colonies 24 days after infection (**e**) or wells with Oct4-eGFP⁺ colonies 8 days after transgene induction (**f**) and treated with the indicated compounds for the days (D) indicated ($n = 2$ independent experiments).



Extended Data Figure 5 | Day 6 continuation analysis on day 16.

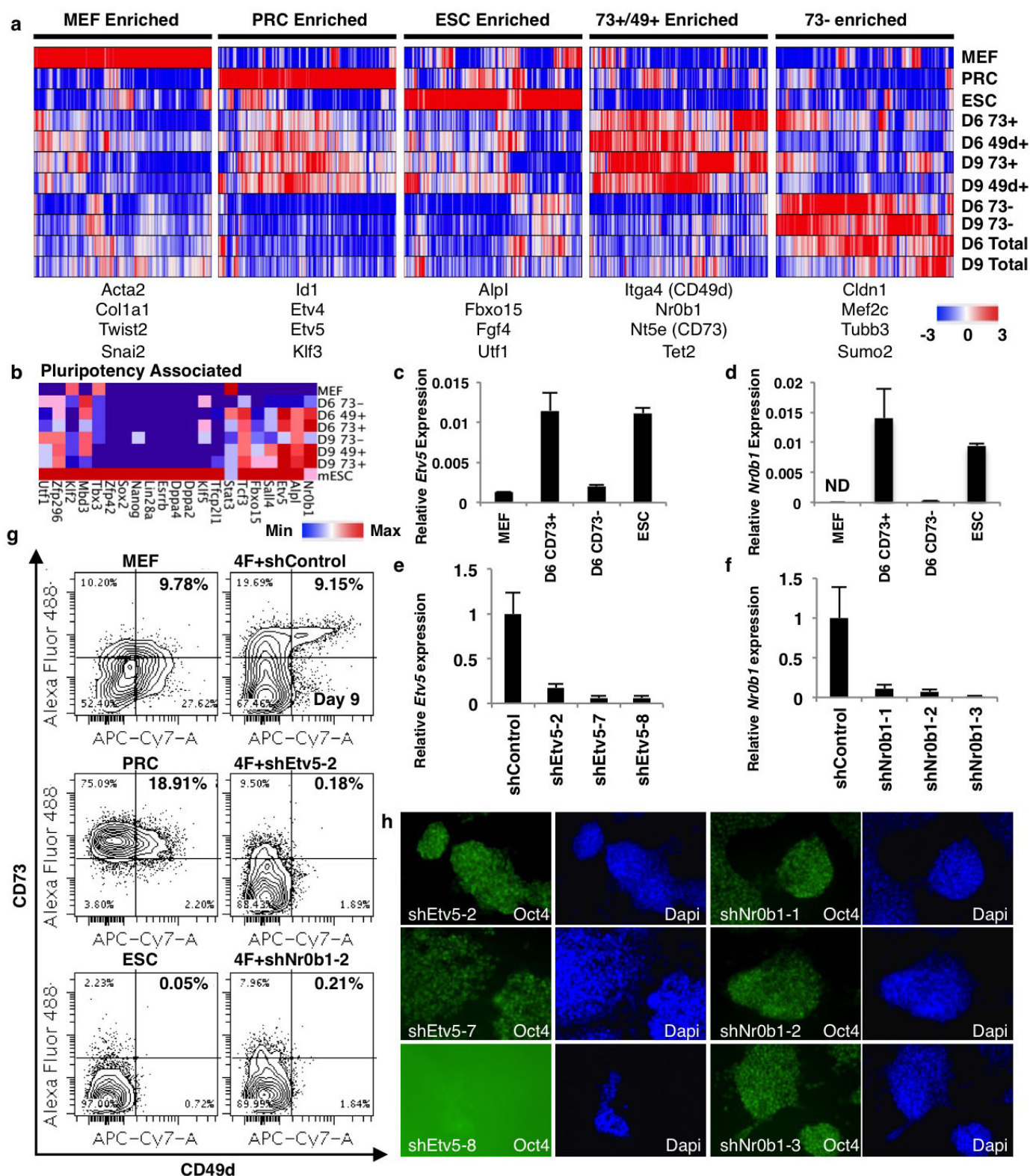
a, Schematic of continuation analysis. Reprogramming populations were sorted for poised (CD73^{high} or CD49d^{high}) and non-poised (CD73^{low}) populations on day 6, cultured for 10 days on a 3 cm plate and analysed by mass cytometry on day 16. **b**, Morphology of CD49d^{high}, CD73^{high} or CD73^{low} cells sorted on day 6 and inspected on day 10. Poised CD49d^{high} or CD73^{high} cells form

compact colonies within several days of sorting while non-poised CD73^{low} cells fail to do so. **c**, SPADE analysis (day 16) of cells sorted at day 6 for CD73^{high}/CD49d^{high}, CD73^{high}, CD49d^{high} and CD73^{low} expression (continued from Fig. 3h). Boxes highlight a SSEA1^{high} CD326^{high} branch that is unique to the poised populations. Colours bars (bottom) ArcSinh-transformed counts for each marker.



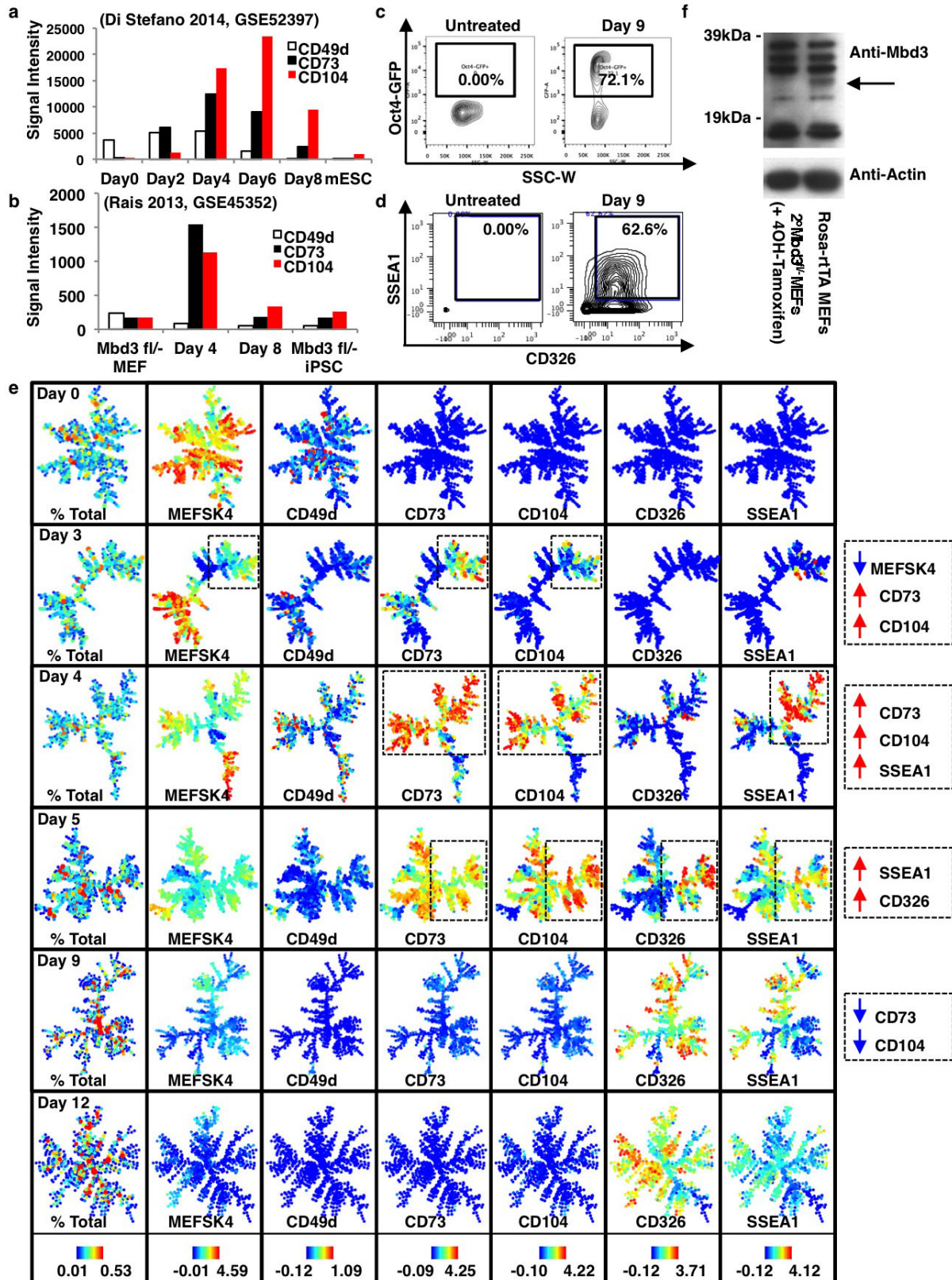
Extended Data Figure 6 | Continuation analysis replicates confirm a SSEA1^{high} CD326^{high} branch that is unique to poised populations. Continuation analysis replicates for reprogramming-prone (CD73^{high}/CD49d^{high}, CD73^{high}, CD49d^{high}) and non-prone (CD73^{low}) populations.

Boxes highlight a SSEA1^{high} CD326^{high} branch that is unique to the poised populations. Colours bars (bottom) represent absolute percentages (left panel) and ArcSinh-transformed counts for each marker.



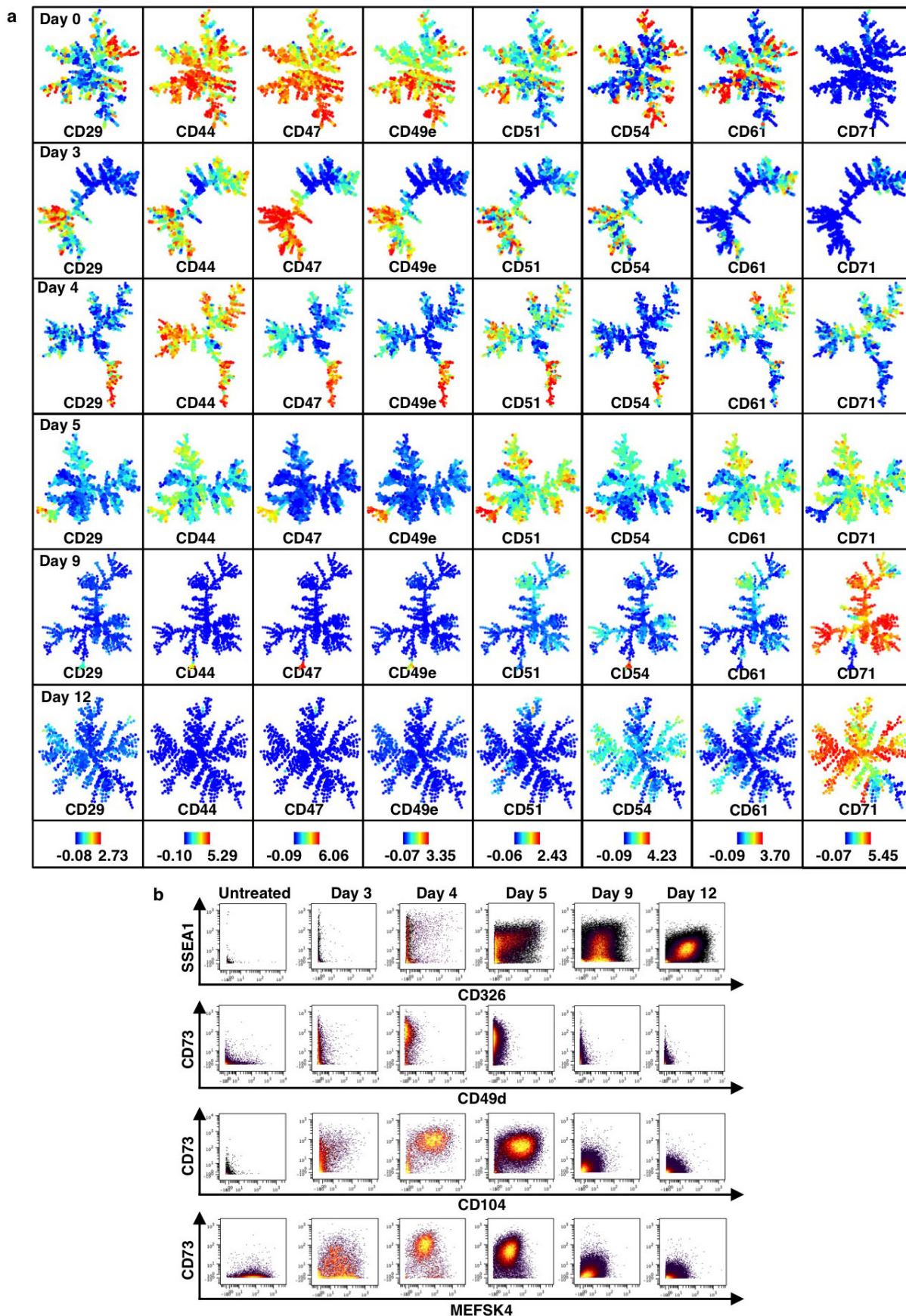
Extended Data Figure 7 | Molecular characterization of reprogramming-prone intermediates. **a**, Genes differentially expressed between reprogramming-prone (day 6 or day 9 CD73^{high} or CD49d^{high}) and non-prone (CD73^{low}) populations. Genes with more than twofold differential expression between reprogramming-prone and non-prone were selected and *k*-means clustered (*k* = 5) with control and total reprogramming population expression values. **b**, Heat map of pluripotency-associated genes shown in Fig. 4a (log₂).

c, d, Quantitative PCR verification of *Etv5* (**c**) and *Nr0b1* (**d**) expression levels (*n* = 3 technical replicates). **e, f**, *Etv5* (**e**) and *Nr0b1* (**f**) knockdown qPCRs (*n* = 3 technical replicates). **g**, Representative FACS plots for day 9 CD73^{high}/CD49d^{high} quantification shown in Fig. 4b. **h**, Demonstration of ESC self-renewal after infection with *Etv5* and *Nr0b1* hairpins. All infected ESCs continue to express *Oct4* after passaging except ESCs infected with sh*Etv5*-8 (*n* = 1).

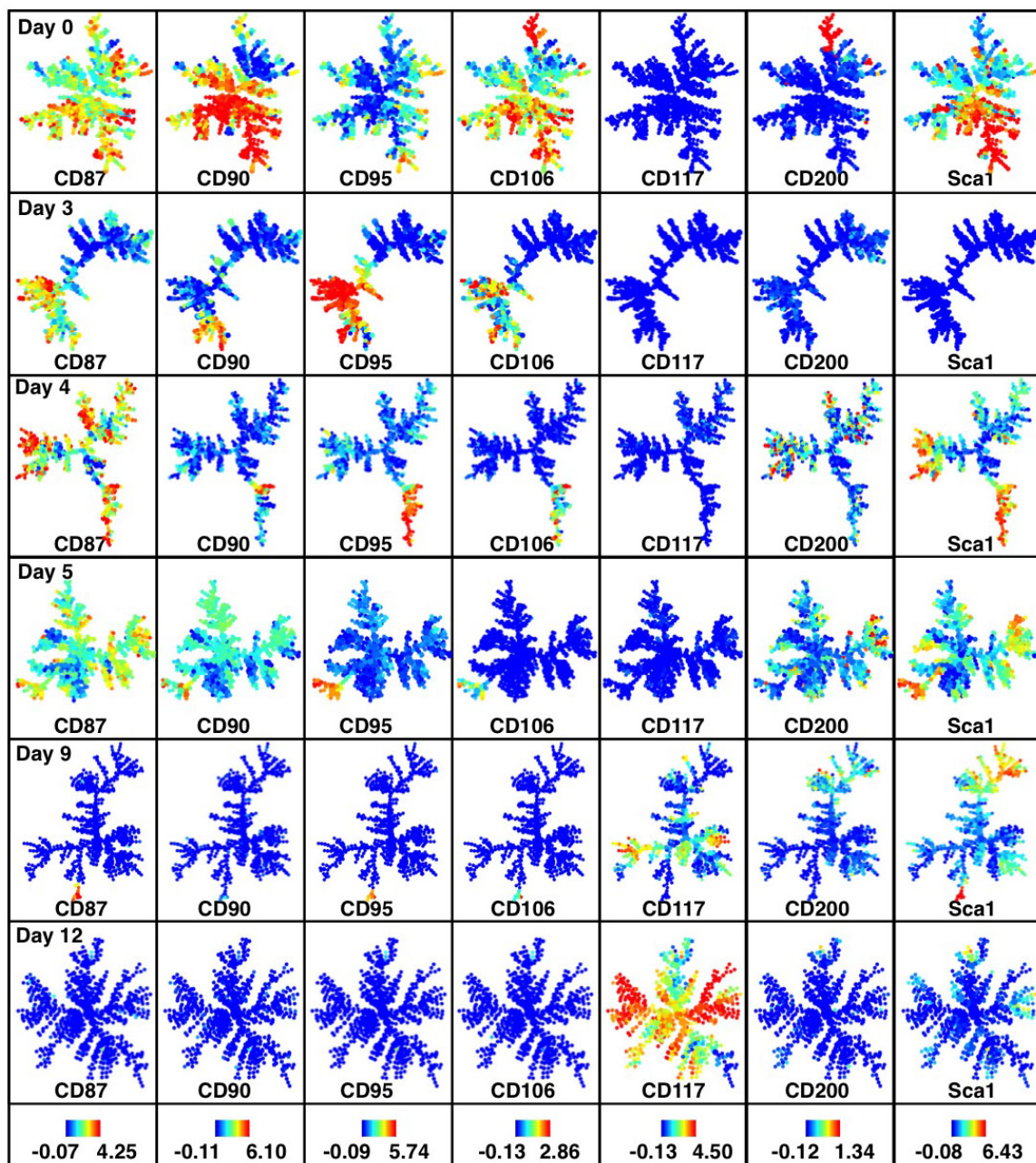


Extended Data Figure 8 | Characterization of high-efficiency reprogramming systems. **a**, **b**, Expression analysis for CD49d, CD73 and CD104 for previously reported highly efficient reprogramming systems generated by transient expression of C/EBP α ¹³ (**a**) or *Mbd3* depletion¹² (**b**). **c**, *Oct4*-GFP transgene reporter signal and **d**, SSEA1 and CD326 levels for the *Mbd3*^{fl/-} secondary reprogramming MEFs for untreated (left) and 9 days after induction (right). **e**, SPADE analysis for reprogramming *Mbd3*^{fl/-} secondary MEFs at days 0, 3, 6, 9 and 12 using all surface markers by mass cytometry.

Percentage totals of cells and representative markers are shown for each time point. Remaining markers are shown in Extended Data Figs 9 and 10. Colours bars represent absolute percentages (left) and ArcSinh-transformed counts for each marker. **f**, Verification of Mbd3 loss in passage 3 *Rosa26-CreER*, *Mbd3*^{fl/-} secondary MEFs after treatment with 4OH-tamoxifen. Mbd3 levels were compared with passage 3 *Rosa-rtTA*⁺ MEFs. While there are several unspecific bands, there is clearly one band around the expected size of Mbd3 absent in 4OH-tamoxifen-treated cells (arrow).



Extended Data Figure 9 | SPADE and biaxial analysis for secondary *Mbd3*^{fl/-} MEFs. **a**, SPADE analysis for 2° *Mbd3*^{fl/-} MEF reprogramming populations (continued from Extended Data Fig. 8e). Colour bars (bottom) represent ArcSinh-transformed counts for each marker. **b**, Biaxial plots for selected markers.



Extended Data Figure 10 | SPADE analysis for secondary *Mbd3*^{fl/-} MEFs. SPADE analysis for 2° *Mbd3*^{fl/-} reprogramming populations (continued from Extended Data Fig. 8e). Colours bars (bottom) represent ArcSinh-transformed counts for each marker.

Signalling thresholds and negative B-cell selection in acute lymphoblastic leukaemia

Zhengshan Chen^{1*}, Seyedmehdi Shojaei^{1*}, Maïke Buchner¹, Huimin Geng¹, Jae Woong Lee¹, Lars Klemm¹, Björn Titz², Thomas G. Graeber², Eugene Park¹, Ying Xim Tan³, Anne Satterthwaite⁴, Elisabeth Paietta⁵, Stephen P. Hunger⁶, Cheryl L. Willman⁷, Ari Melnick⁸, Mignon L. Loh⁹, Jae U. Jung¹⁰, John E. Coligan¹¹, Silvia Bolland¹², Tak W. Mak¹³, Andre Limnander¹⁴, Hassan Jumaa¹⁵, Michael Reth¹⁶, Arthur Weiss³, Clifford A. Lowell¹ & Markus Müschen¹

B cells are selected for an intermediate level of B-cell antigen receptor (BCR) signalling strength: attenuation below minimum (for example, non-functional BCR)¹ or hyperactivation above maximum (for example, self-reactive BCR)^{2,3} thresholds of signalling strength causes negative selection. In ~25% of cases, acute lymphoblastic leukaemia (ALL) cells carry the oncogenic *BCR-ABL1* tyrosine kinase (Philadelphia chromosome positive), which mimics constitutively active pre-BCR signalling^{4,5}. Current therapeutic approaches are largely focused on the development of more potent tyrosine kinase inhibitors to suppress oncogenic signalling below a minimum threshold for survival⁶. We tested the hypothesis that targeted hyperactivation—above a maximum threshold—will engage a deletional checkpoint for removal of self-reactive B cells and selectively kill ALL cells. Here we find, by testing various components of proximal pre-BCR signalling in mouse BCR-ABL1 cells, that an incremental increase of Syk tyrosine kinase activity was required and sufficient to induce cell death. Hyperactive Syk was functionally equivalent to acute activation of a self-reactive BCR on ALL cells. Despite oncogenic transformation, this basic mechanism of negative selection was still functional in ALL cells. Unlike normal pre-B cells, patient-derived ALL cells express the inhibitory receptors PECAM1, CD300A and LAIR1 at high levels. Genetic studies revealed that *Pecam1*, *Cd300a* and *Lair1* are critical to calibrate oncogenic signalling strength through recruitment of the inhibitory phosphatases *Ptpn6* (ref. 7) and *Inpp5d* (ref. 8). Using a novel small-molecule inhibitor of INPP5D (also known as SHIP1)⁹, we demonstrated that pharmacological hyperactivation of SYK and engagement of negative B-cell selection represents a promising new strategy to overcome drug resistance in human ALL.

ALL represents the most frequent type of cancer in children and is frequent in adults as well. Although outcomes for patients with ALL have greatly improved over the past four decades, ALL driven by oncogenic tyrosine kinases (*BCR-ABL1* in adults and other oncogenic fusion tyrosine kinases in childhood ALL)¹⁰ remains a clinical problem. Current efforts to improve treatment options are largely focused on the development of more potent tyrosine kinase inhibitors (TKIs). However, responses to TKIs are often short lived. Our group recently identified upregulation of the *BCL6* proto-oncogene in response to TKI treatment as a major mechanism of drug resistance in Philadelphia chromosome positive (Ph⁺) ALL¹¹. Here, we propose a strategy to overcome drug

resistance in ALL based on the pharmacological hyperactivation of SYK.

Pre-BCR signals are initiated from immunoreceptor tyrosine-based activation motifs (ITAMs) in the cytoplasmic tail of immunoglobulin (Ig) α (*CD79A*) and Ig β (*CD79B*) signalling chains¹², and are essential for survival and proliferation of normal pre-B cells. However, hyperactive signalling from a self-reactive pre-BCR, owing to the ubiquitous presence of self-antigen, induces negative selection and cell death³. Here, we observed that Ph⁺ ALL cells consistently lack surface expression of ITAM-bearing Ig α and Ig β signalling chains (Extended Data Fig. 1a). Seemingly in contrast to compromised Ig α and Ig β expression, multiple components of proximal pre-BCR signalling were activated downstream of the BCR-ABL1 tyrosine kinase (Fig. 1a). These findings demonstrate that oncogenic BCR-ABL1 supplants ITAM-dependent signalling and mimics a constitutively active pre-BCR through engagement with its proximal signalling cascade. Besides BCR-ABL1, mimicry of BCR signalling was previously demonstrated for a number of viral oncoproteins including Epstein-Barr virus latent membrane protein 2A (LMP2A)¹³.

Reconstitution of Ig α expression induced strong tyrosine phosphorylation of proximal pre-BCR signalling molecules, followed by cell death (Extended Data Fig. 1b–d). Likewise, Ph⁺ ALL cells from three patients were highly sensitive to reactivation of ITAM-dependent signalling (LMP2A¹³; Extended Data Fig. 1e, f). Interestingly, activation of ITAM signalling was toxic in leukaemic but not in normal pre-B cells (Extended Data Fig. 1b). We therefore tested whether BCR-ABL1- and ITAM-dependent activation of proximal pre-BCR signalling are mutually exclusive because both engage the same pre-BCR-associated tyrosine kinases. Consequently, we repeated activation of ITAM signalling in the presence and absence of TKI treatment (imatinib; Fig. 1b). Seemingly paradoxically, treatment with imatinib, although designed to kill leukaemia cells, rescued *BCR-ABL1* ALL cells in this experimental setting, and subsequent washout of imatinib reversed the protective effect (Fig. 1b).

To pinpoint which aspect of proximal pre-BCR signalling is toxic to Ph⁺ ALL cells, we used genetic systems for hyperactivation of Syk, Src and Btk. In contrast to Src and Btk, constitutively active Syk (Syk^{Myr}) induced rapid cell death (Fig. 1b and Extended Data Fig. 2a–d). Hyperactive Syk was synthetically lethal in combination with oncogenic BCR-ABL1, and cytotoxic effects were mitigated by TKI treatment (imatinib; Fig. 1b). Like BCR-ABL1, SYK kinase activity alone mimicked

¹Department of Laboratory Medicine, University of California, San Francisco, California 94143, USA. ²Crump Institute for Molecular Imaging, Department of Molecular and Medical Pharmacology, University of California, Los Angeles, California 90095, USA. ³Rosalind Russell-Ephraim P. Engleman Medical Research Center for Arthritis, Division of Rheumatology, Department of Medicine, Howard Hughes Medical Institute, University of California, San Francisco, California 94143, USA. ⁴Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ⁵Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10466, USA. ⁶Division of Pediatric Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, Philadelphia 19104, USA. ⁷University of New Mexico Cancer Center, Albuquerque, New Mexico 87102, USA. ⁸Departments of Medicine and Pharmacology, Weill Cornell Medical College, New York, New York 10065, USA. ⁹Pediatric Hematology-Oncology, University of California, San Francisco, California 94143, USA. ¹⁰Department of Molecular Microbiology and Immunology, University of Southern California, Los Angeles, California 90033, USA. ¹¹Receptor Cell Biology Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland 20852, USA. ¹²Autoimmunity and Functional Genomics Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland 20852, USA. ¹³The Campbell Family Institute for Breast Cancer Research, University Health Network, 620 University Avenue, Toronto, Ontario M5G 2M9, Canada. ¹⁴Department of Anatomy, University of California, San Francisco, California 94143, USA. ¹⁵Institute of Immunology, University Clinics Ulm, 89081 Ulm, Germany. ¹⁶BIOSS Centre for Biological Signalling Studies and Faculty of Biology, Albert-Ludwigs-Universität Freiburg, and MPI of Immunobiologie and Epigenetics, 79104 Freiburg, Germany.

*These authors contributed equally to this work.

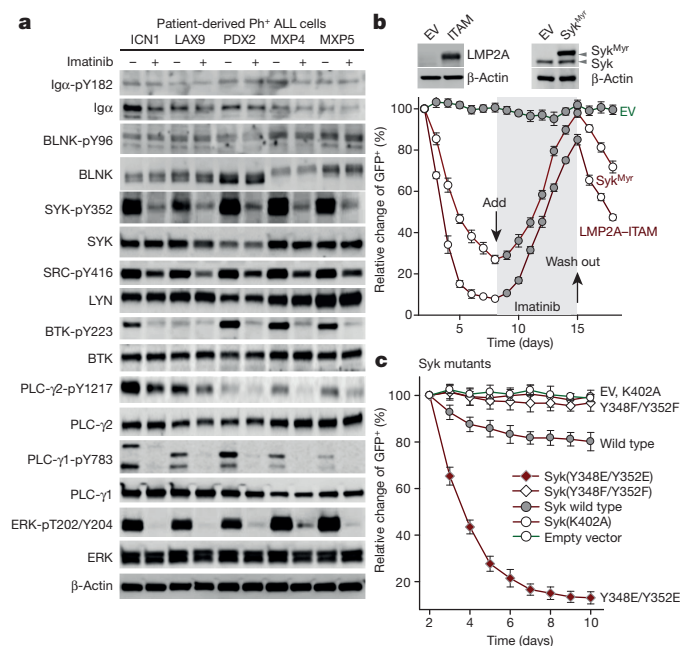


Figure 1 | Reconstitution of defective ITAM signalling induces cell death in BCR-ABL1 ALL cells. **a**, Patient-derived Ph⁺ ALL cells were treated with or without imatinib (10 μmol l⁻¹) for 6 h and phosphorylation (p) of Igα, BLNK, SYK, SRC, BTK, PLC-γ2, PLC-γ1 and ERK1/2 was measured by western blot (*n* = 5). ICN1, LAX9, PDX2, MXP4 and MXP5 are names of patient-derived Ph⁺ ALL cells (see Supplementary Table 1). **b**, BCR-ABL1 ALL cells transduced with GFP-tagged LMP2A-ITAM, Syk^{Myr} or an empty vector (EV) were monitored over time in the presence or absence of 0.5 μmol l⁻¹ imatinib by flow cytometry. The expression level of LMP2A and Syk^{Myr} were measured by western blot. **c**, BCR-ABL1 ALL cells were transduced with GFP-tagged wild-type Syk or Syk mutant vectors (Y348E/Y352E, Y348F/Y352F, K402A) or an empty vector and relative changes of transduced (GFP⁺) cells were monitored by flow cytometry. Data are presented as means ± standard deviation (s.d.) from three independent experiments (**b**, **c**).

constitutively active pre-BCR signalling and was sufficient to transform mouse pro-B cells (Extended Data Fig. 2e). Interestingly, BCR-ABL1 kinase activity induced phosphorylation of SYK at interdomain B (Fig. 1a), which relieves the autoinhibitory conformation of Syk¹⁴. To study the specific function of Syk interdomain B (Y348 and Y352) tyrosines in BCR-ABL1 ALL cells, we tested loss (Y→F) and phosphomimetic gain (Y→E) of function mutants of Syk. Empty vectors, kinase-dead Syk(K402A) and wild-type Syk were used as controls (Fig. 1c). In the absence of constitutive membrane localization, wild-type Syk had only minor toxic effects on ALL cells. Interestingly, however, the expression of Syk carrying phosphomimetic mutations of interdomain B tyrosines (Y348/Y352→E348/E352) induced rapid cell death (Fig. 1c). These findings highlight the relevance of Syk interdomain B tyrosines and suggest that pharmacological approaches to increase tyrosine phosphorylation of Syk interdomain B may be useful to kill Ph⁺ ALL cells. To study whether Syk tyrosine kinase activity is required for induction of cell death in pre-B ALL cells, we used a Syk tyrosine kinase inhibitor, PRT062607 (PRT). Transduction with constitutively active Syk^{Myr} induced rapid cell death, which was rescued by pre-treatment with PRT 1 day before transduction with Syk^{Myr}. Interestingly, a 1-day lapse of PRT treatment and transient hyperactivation of Syk was sufficient to commit pre-B ALL cells to cell death (Extended Data Fig. 1g).

In the absence of direct strategies for Syk hyperactivation, we studied pharmacological inhibition of negative regulators of Syk. In normal pre-B cells, activation of Syk downstream of the pre-BCR is negatively regulated by inhibitory surface receptors that bear immunoreceptor tyrosine-based inhibitory (ITIM)¹⁵ motifs in their cytoplasmic tail. A systematic screen identified 109 ITIM-bearing receptors in the human

genome¹⁶, 62 of which are expressed in B cells. Compared to normal pre-B cells and mature B-cell lymphoma, the majority of ITIM receptors were upregulated in Ph⁺ ALL cells. On the basis of the ratio of expression values in Ph⁺ ALL compared to pre-B cells and mature B-cell lymphoma, PECAM1, CD300A and LAIR1 were identified as among the top-ranking ITIM receptors, which was confirmed by flow cytometry (Extended Data Fig. 3).

To determine whether high expression levels of ITIM-bearing receptors influence the course of human ALL, we segregated patients from two clinical trials (the Children's Oncology Group (COG) P9906 study and the Eastern Cooperative Oncology Group (ECOG) 2993 study) into two groups on the basis of whether they had higher or lower than median expression levels of PECAM1, CD300A and LAIR1 at the time of diagnosis. Higher than median expression levels of ITIM receptors on ALL cells at the time of diagnosis predicted shorter overall and relapse-free survival (Extended Data Fig. 4a–e). These findings identify ITIM-bearing inhibitory receptors as a novel biomarker with potential use in risk stratification of children and adults with ALL.

To measure the functional consequences of ITIM-receptor deletion, pre-B cells from the bone marrow of *Pecam1*^{-/-} and *Cd300a*^{-/-}, as well as *Lair1*^{fl/fl} mice and wild-type controls were propagated in the presence of interleukin (IL)-7 or transformed with BCR-ABL1 to model human Ph⁺ ALL. *Lair1*^{fl/fl} ALL cells were retrovirally transduced with 4-hydroxytamoxifen (4-OHT)-inducible Cre. Loss of ITIM receptors had no significant effects on the proliferation and survival of normal pre-B cells (Extended Data Fig. 5a). In contrast, in the absence of ITIM-bearing receptors, pre-B ALL cells underwent cellular senescence and cell cycle arrest and failed to form colonies (Fig. 2a and Extended Data Fig. 5a, b) in parallel with activation of cell cycle checkpoint molecules and increased levels of cytoplasmic reactive oxygen species (ROS; Extended Data Fig. 4g, h). Importantly, inducible Cre-mediated ablation of *Lair1* surface expression (Extended Data Fig. 4f) resulted in massive hyperactivation of Syk (Y352), Src kinases (Y416) and Erk (T202/Y204; Fig. 2b), which promotes negative selection of autoreactive B-cell clones during early B-cell development¹⁷. In agreement with these findings, Cre-mediated deletion of *Lair1* caused rapid cell death *in vitro*, remission of leukaemia *in vivo* and significantly prolonged survival of transplant-recipient mice (*P* = 0.0003, log-rank test; Fig. 2c, d and Extended Data Fig. 5c).

The surface receptors PECAM1, CD300A and LAIR1 attenuate pre-BCR signalling through ITIM-dependent recruitment and activation of inhibitory phosphatases (for example, PTPN6 (also known as SHP1), INPP5D)^{7,8}. For this reason, we performed experiments to determine whether *Lair1* contributes to activation of *Ptpn6* and *Inpp5d*. Consistent with a role of *Lair1* in the recruitment and activation of Ptpn6 and Inpp5d, activating tyrosine phosphorylation of Ptpn6 (Y564) and Inpp5d (Y1020) was reduced by three- to fourfold upon inducible deletion of *Lair1* (Extended Data Fig. 5d). In genetic rescue experiments, we demonstrated that intact ITIM motifs in the cytoplasmic tails of *Pecam1*, *Lair1* and *Cd300a* are critical for the survival of pre-B ALL cells: *Pecam1*^{-/-}, *Lair1*^{-/-} and *Cd300a*^{-/-} pre-B cells were transduced with green fluorescent protein (GFP)-tagged vectors for reconstitution of *Pecam1*, *Lair1* and *Cd300a* bearing either wild-type or mutant (Y→F/A) ITIM motifs or GFP empty vector controls, and then transformed by BCR-ABL1 (Fig. 2e–g). Reconstitution with wild-type-ITIM *Pecam1*, *Lair1* and *Cd300a* rescued survival and proliferation, whereas reconstitution with receptors carrying tyrosine-mutant ITIMs had no effect (Fig. 2e–g).

The phosphatases PTPN6 (ref. 7), INPP5D⁸ and PTPN11 (also known as SHP2)¹⁸ can all bind to ITIM motifs. We determined their mechanistic contribution to calibration of oncogenic signalling in a genetic rescue experiment: *Lair1*^{fl/fl} ALL cells were transduced with GFP-tagged expression vectors of constitutively active or phosphatase-inactive forms of Ptpn6, Inpp5d and Ptpn11 (Fig. 3a and Extended Data Fig. 5e). Expression of constitutively active *Inpp5d* or *Ptpn6*, but not *Ptpn11*, rescued cell death after Cre-mediated deletion of *Lair1*. Interestingly,

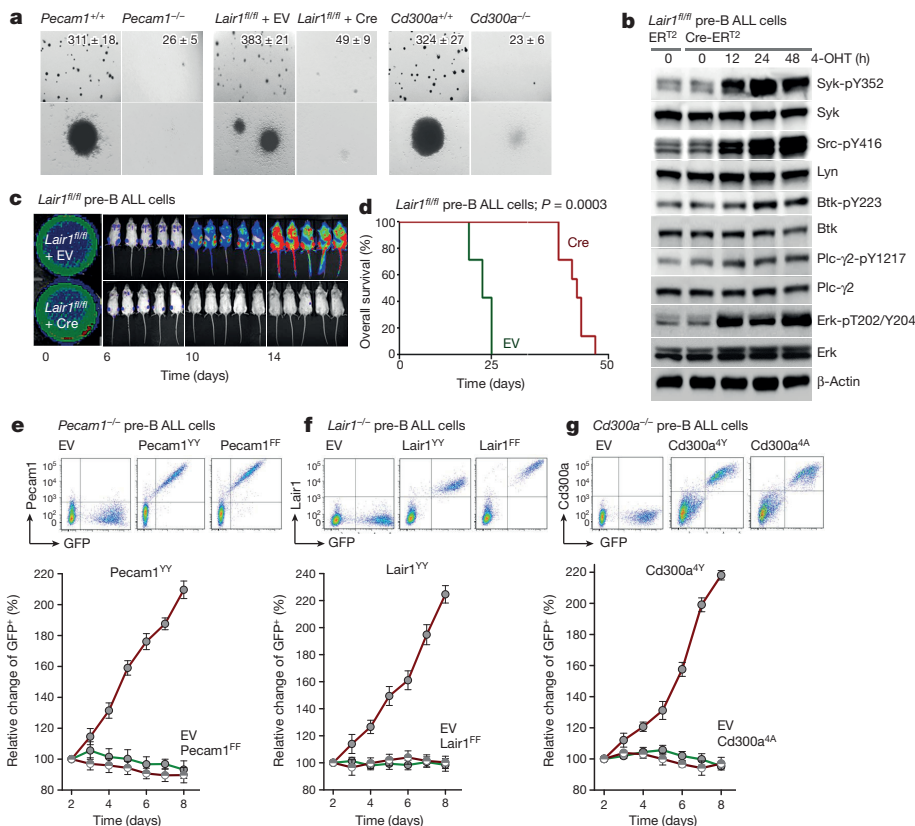


Figure 2 | Inhibitory ITIM-bearing receptors are critical for pre-B leukaemogenesis. **a**, Pre-B cells from *Pecam1*^{-/-}, *Cd300a*^{-/-} and *Lair1*^{fl/fl} mice and wild-type controls were propagated with IL-7 and transduced with BCR-ABL1. *Lair1*^{fl/fl} ALL cells were transduced with 4-OHT-inducible Cre. Colony formation assays were performed, showing photomicrographs of colonies at ×1 (top) and ×10 (bottom) magnification. Numbers at top right indicate mean colony number ± s.d. **b**, Effects of inducible deletion of *Lair1* on phosphorylation (p) levels of Syk, Src, Btk, Plc-γ2 and Erk were measured by western blot. Data are representative of three independent experiments. ER, oestrogen receptor. **c**, *Lair1*^{fl/fl} ALL cells were labelled with firefly luciferase, transduced with 4-OHT-inducible Cre or empty vector control (EV), treated with 4-OHT for 24 h and transplanted into sublethally irradiated NOD/SCID mice, and leukaemia burden was measured by luciferase bioimaging. **d**, A Kaplan-Meier analysis compared overall survival of transplant recipients in the two groups (*n* = 7 for each group). *P* value was calculated by log-rank test. **e–g**, *Pecam1*^{-/-}, *Lair1*^{-/-} (4-OHT-induced deletion) and *Cd300a*^{-/-} pre-B cells were reconstituted with wild-type (*Pecam1*^{YY}, *Lair1*^{YY}, *Cd300a*^{4Y}), *Pecam1*^{FF} (Y679F/Y702F), *Lair1*^{FF} (Y228F/Y257F), *Cd300a*^{4A} (Y231A/Y255A/Y267A/Y293A) mutant vectors or empty vector, and then transformed by BCR-ABL1. The expression levels of wild-type or mutant receptors were monitored by flow cytometry. Data are presented as means ± s.d. from three independent experiments (**a**, **e–g**).

inducible deletion of *Ptpn6* or *Inpp5d* was sufficient to cause cell death and a sharp increase of cellular ROS levels in ALL cells (Fig. 3b, c and Extended Data Figs 6b–e, 7a). Given that phosphatases are sensitive to reversible inactivation by cysteine oxidation of their active sites¹⁹, we tested whether deletion of one single phosphatase triggers an ROS-mediated chain reaction of phosphatase inactivation. Using antibodies against phosphatases in inactivated oxidized conformation, we found that deletion of either *Ptpn6* or *Inpp5d* caused widespread cysteine oxidation and inactivation of multiple other phosphatases (Extended Data Fig. 7b). Inducible ablation of *Ptpn6* or *Inpp5d* caused increased expression of Arf and p53 cell cycle checkpoint molecules, G₀/1 cell cycle arrest and 15- to 40-fold reduced colony formation capacity (Fig. 3d, e and Extended Data Fig. 7c–e). In an *in vivo* transplant experiment, inducible *in vivo* deletion of *Ptpn6* or *Inpp5d* significantly reduced penetrance and extended the latency of the leukaemia (Fig. 3f; *P* < 0.0005, log-rank test). These findings reveal a novel and unexpected vulnerability and suggest that ITIM-bearing receptors and inhibitory phosphatases represent a novel class of therapeutic targets in pre-B ALL. Both PTPN6 and INPP5D attenuate ITAM-dependent pre-BCR signalling in normal pre-B cells^{7,8}. Cre-mediated depletion of *Ptpn6* or *Inpp5d* protein resulted in strong hyperactivation of Syk (Y352; Fig. 3g, h). While PTPN6 directly dephosphorylates ITAMs and SYK⁷, INPP5D hydrolyses the membrane anchor PIP3 and thereby inhibits formation and maintenance of ITAM-dependent signalling complexes at the cell membrane²⁰. Pre-treatment with PRT largely rescued cell death, demonstrating that hyperactivation of Syk is a mechanistic requirement for induction of cell death (Fig. 3i, j).

B-lineage Ph⁺ ALL and myeloid-lineage chronic myeloid leukaemia (CML) are both driven by BCR-ABL1. As opposed to Ph⁺ ALL, however, defective expression of ITIM receptors, *Ptpn6* or *Inpp5d* had no functional consequences in a mouse model for CML (Extended Data Figs 8 and 9). Consistent with these findings, PTPN6 and INPP5D are highly expressed in patient-derived Ph⁺ ALL (*n* = 5) but barely detectable

in CML cells (*n* = 5; Extended Data Fig. 6a). To test whether B-cell-inherent mechanisms of negative selection are still active and examine the underlying reason for the divergent behaviour of B-lineage and myeloid leukaemia, we engineered B-cell-lineage Ph⁺ ALL cells with a doxycycline-inducible vector system for expression of *Cebpa*²¹, which results in myeloid-lineage reprogramming (Extended Data Fig. 10a, b). BCR-ABL1-transformed pre-B ALL cells were transduced with GFP-tagged Cre and reprogrammed into myeloid-lineage leukaemia cells. While inducible ablation of *Lair1*, *Ptpn6* or *Inpp5d* resulted in rapid cell death among B-lineage (CD19⁺ B220⁺ Mac1⁻) ALL cells, myeloid-lineage reprogramming (CD19⁻ B220⁻ Mac1⁺) rendered leukaemia cells resistant to the effects of inducible deletion (Extended Data Fig. 10c–e). These findings support a scenario in which Ph⁺ ALL cells are subject to B-cell-specific negative selection against hyperactive Syk tyrosine kinase signalling emanating from a self-reactive BCR, or its oncogenic mimic BCR-ABL1. Inducible expression of *Cebpa* subverts B-cell lineage commitment and raises the threshold for tyrosine kinase hyperactivation to trigger cell death. In this context, it is interesting to note that multiple genetic lesions in human pre-B ALL target transcription factors that mediate B-cell lineage commitment, including *IKZF1*, *PAX5* and *EBF1* (ref. 22). Although their mechanistic role is not known, we propose that deletions of *IKZF1*, *PAX5* and *EBF1*, like downregulation of *PAX5* in the context of *Cebpa* expression, reduce the stringency of negative selection against hyperactive tyrosine kinase signalling.

A small-molecule inhibitor against INPP5D, 3-α-aminocholestane (3AC)⁹ (Extended Data Fig. 10f) selectively inhibited enzymatic activity of INPP5D (half-maximum inhibitory concentration (IC₅₀) ~ 2.5 μmol l⁻¹) but not the related phosphatases INPPL1 (also known as SHIP2) and PTEN (IC₅₀ > 20 μmol l⁻¹)⁹. Treatment of patient-derived Ph⁺ ALL cells with 3AC induced strong hyperactivation of SYK (Fig. 4a). In patient-derived myeloid CML samples, baseline levels of SYK activity were very low and not responsive to 3AC treatment (Extended Data Fig. 10g). Biochemical characterization of 3AC-mediated inhibition of

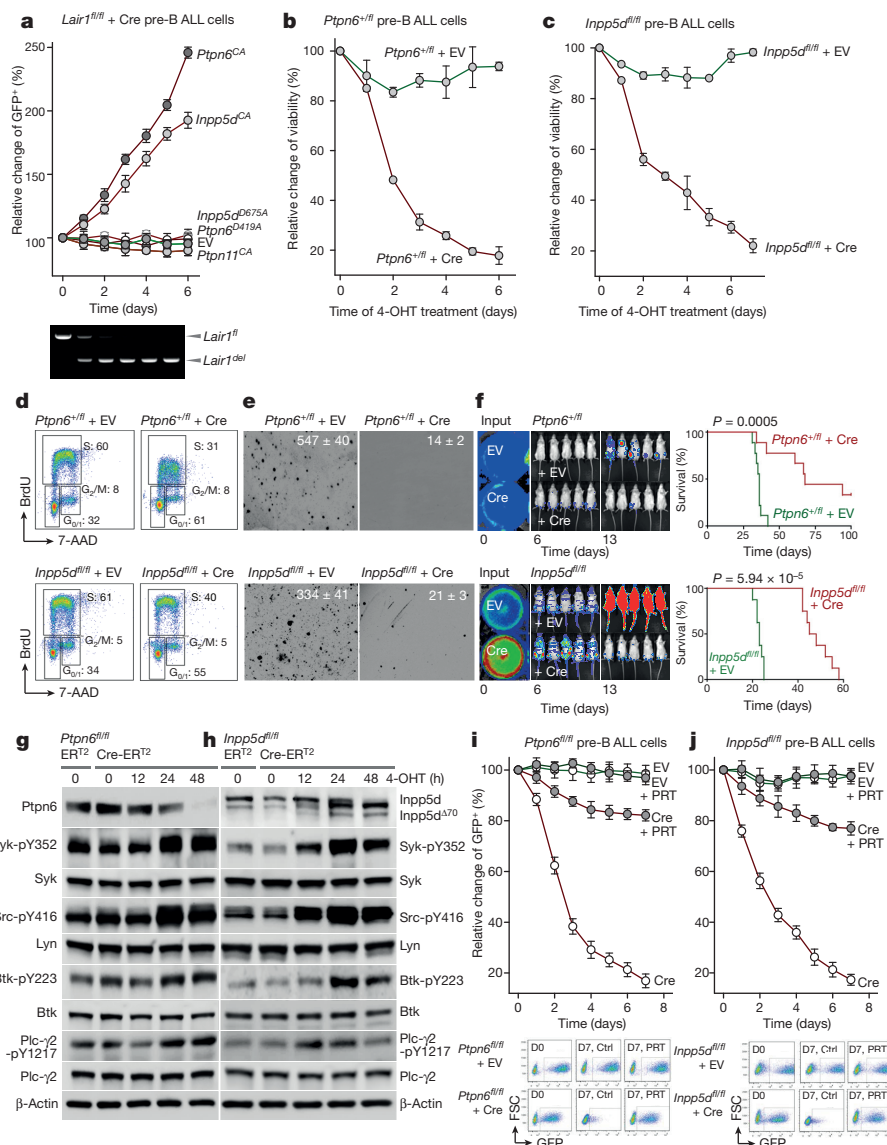


Figure 3 | ITIM-dependent activation of Ptpn6 and Inpp5d phosphatases enables pre-B leukaemogenesis. **a**, *Lair1^{fl/fl}* ALL cells were transduced with 4-OHT-inducible Cre. After antibiotic selection, leukaemia cells were transduced with GFP-tagged empty vector (EV) or overexpression vectors for constitutively active (CA) forms of *Ptpn6* (SH2 domain deleted), *Inpp5d* (CD8-*Inpp5d*), *Ptpn11* (*Ptpn11^{D61A}*), or phosphatase-inactive mutants (*Ptpn6^{D419A}*, *Inpp5d^{D675A}*). After addition of 4-OHT, Cre-mediated deletion of *Lair1* (*Lair1^{del}*) was monitored by polymerase chain reaction (PCR). Percentages of GFP⁺ cells were measured by flow cytometry. **b**, **c**, Inducible activation of Cre in *Ptpn6^{+/fl}* (**b**) and *Inpp5d^{fl/fl}* (**c**) BCR-*ABL1*-transformed ALL cells resulted in depletion of transduced cells. **d**, **e**, Effects of deletion of *Ptpn6* and *Inpp5d* on proliferation (cell cycle analysis, BrdU) (**d**) and colony formation ability (**e**) were measured. **d**, Numbers indicate percentage of cells in each cell cycle phase. **e**, Numbers at top right indicate mean colony number

INPP5D in patient-derived Ph⁺ ALL cells revealed potent and transient hyperactivation of proximal pre-BCR signalling molecules (Fig. 4a). Treatment of patient-derived TKI-resistant Ph⁺ ALL cells with 3AC induced cell death within 4 days. Importantly, pre-treatment of Ph⁺ ALL cells with PRT largely protected Ph⁺ ALL cells against 3AC-induced cell death (Fig. 4b), demonstrating that hyperactivation of SYK is required for induction of cell death. Dose-response analyses revealed that 3AC is selectively toxic for patient-derived Ph⁺ ALL cells ($IC_{50} = 2.8 \mu\text{mol l}^{-1}$; $n = 5$) compared to mature B-cell lymphoma ($n = 5$; Extended Data Fig. 10h). We next studied drug responses in a panel of six

cases of Ph⁺ ALL from patients who relapsed under TKI therapy, including three cases with global TKI resistance owing to the BCR-*ABL1*(T315I) mutation. As expected, treatment with imatinib had no effect in BCR-*ABL1*(T315I) cases (Extended Data Fig. 10i). In contrast, 3AC induced massive cell death (>95%) in all six cases of Ph⁺ ALL regardless of BCR-*ABL1* mutation status (Extended Data Fig. 10i). Likewise, treatment of NOD/SCID transplant-recipient mice carrying TKI-resistant patient-derived (BCR-*ABL1*(T315I)) Ph⁺ ALL cells with 3AC significantly prolonged overall survival ($P = 0.0002$, log-rank test; Fig. 4c) and reduced leukaemia burden (Fig. 4d). While further studies are needed

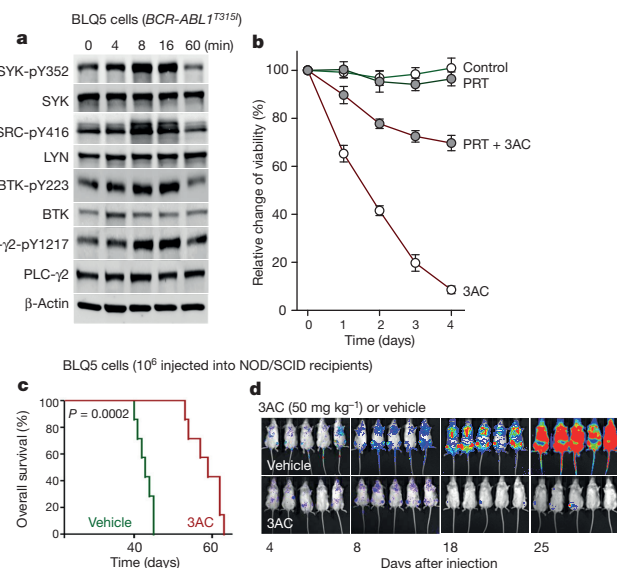


Figure 4 | Small-molecule inhibition of INPP5D induces hyperactivation of SYK and triggers a deletional checkpoint in pre-B ALL cells. **a**, Patient-derived Ph⁺ ALL cells (BLQ5) were treated with 3AC (10 μmol l⁻¹) for the times indicated and phosphorylation of SYK, SRC, BTK and PLC-γ2 were measured by western blot. Data are representative of three independent experiments. **b**, ALL cells were treated with vehicle, PRT (2.5 μmol l⁻¹), 3AC (7.5 μmol l⁻¹) alone, or pre-treated with PRT for 2 days, after which 3AC was added. Viability was monitored by flow cytometry. Error bars represent means ± s.d. from three independent experiments. **c**, **d**, TKI-resistant patient-derived Ph⁺ ALL cells (BLQ5) were labelled with firefly luciferase and injected into sublethally irradiated NOD/SCID mice, treated with either 3AC or vehicle (50 mg kg⁻¹, daily intraperitoneal injection, *n* = 7 per group). Overall survival of recipient mice in the two groups was compared by Kaplan–Meier analysis (*P* value calculated by log-rank test) (**c**) and leukaemia burden was measured by luciferase bioluminescence (**d**).

to optimize pharmacological targeting of this pathway, these experiments identify transient hyperactivation of SYK and engagement of negative B-cell selection as a powerful new strategy to overcome drug resistance in Ph⁺ ALL.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 September 2013; accepted 13 January 2015.

Published online 23 March 2015.

- Lam, K. P., Kühn, R. & Rajewsky, K. *In vivo* ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell* **90**, 1073–1083 (1997).
- Nemazee, D. A. & Bürki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* **337**, 562–566 (1989).
- Keenan, R. A. *et al.* Censoring of autoreactive B cell development by the pre-B cell receptor. *Science* **321**, 696–699 (2008).

- Klein, F. *et al.* The BCR-ABL1 kinase bypasses selection for the expression of a pre-B cell receptor in pre-B acute lymphoblastic leukemia cells. *J. Exp. Med.* **199**, 673–685 (2004).
- Feldhahn, N. *et al.* Mimicry of a constitutively active pre-B cell receptor in acute lymphoblastic leukemia cells. *J. Exp. Med.* **201**, 1837–1852 (2005).
- Cortes, J. E. *et al.* Ponatinib in refractory Philadelphia chromosome-positive leukemias. *N. Engl. J. Med.* **367**, 2075–2088 (2012).
- Pao, L. I. *et al.* B cell-specific deletion of protein-tyrosine phosphatase Shp1 promotes B-1a cell development and causes systemic autoimmunity. *Immunity* **27**, 35–48 (2007).
- O'Neill, S. K. *et al.* Monophosphorylation of CD79a and CD79b ITAM motifs initiates a SHIP-1 phosphatase-mediated inhibitory signaling cascade required for B cell anergy. *Immunity* **35**, 746–756 (2011).
- Brooks, R. *et al.* SHIP1 inhibition increases immunoregulatory capacity and triggers apoptosis of hematopoietic cancer cells. *J. Immunol.* **184**, 3582–3589 (2010).
- Roberts, K. G. *et al.* Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N. Engl. J. Med.* **371**, 1005–1015 (2014).
- Duy, C. *et al.* BCL6 enables Ph⁺ acute lymphoblastic leukaemia cells to survive BCR-ABL1 kinase inhibition. *Nature* **473**, 384–388 (2011).
- Reth, M. Antigen receptor tail clue. *Nature* **338**, 383–384 (1989).
- Anderson, L. J. & Longnecker, R. EBV LMP2A provides a surrogate pre-B cell receptor signal through constitutive activation of the ERK/MAPK pathway. *J. Gen. Virol.* **89**, 1563–1568 (2008).
- Kulathu, Y., Grothe, G. & Reth, M. Autoinhibition and adapter function of Syk. *Immunol. Rev.* **232**, 286–299 (2009).
- Bolland, S. & Ravetch, J. V. Inhibitory pathways triggered by ITIM-containing receptors. *Adv. Immunol.* **72**, 149–177 (1999).
- Staub, E., Rosenthal, A. & Hinzmann, B. Systematic identification of immunoreceptor tyrosine-based inhibitory motifs in the human proteome. *Cell. Signal.* **16**, 435–456 (2004).
- Limnander, A. *et al.* STIM1, PKC-δ and RasGRP set a threshold for proapoptotic Erk signaling during B cell development. *Nature Immunol.* **12**, 425–433 (2011).
- Zhang, S. Q. *et al.* Shp2 regulates SRC family kinase activity and Ras/Erk activation by controlling Csk recruitment. *Mol. Cell* **13**, 341–355 (2004).
- Meng, T.-C., Fukada, T. & Tonks, N. K. Reversible oxidation and inactivation of protein tyrosine phosphatases *in vivo*. *Mol. Cell* **9**, 387–399 (2002).
- Bolland, S., Pearce, R. N., Kurosaki, T. & Ravetch, J. V. SHIP modulates immune receptor responses by regulating membrane association of Btk. *Immunity* **8**, 509–516 (1998).
- Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663–676 (2004).
- Mullighan, C. G. *et al.* BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* **453**, 110–114 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. W. Hendriks for encouragement and critical discussions, A. Park and C. Lin and all the members of the Müschen laboratory, L. Tian and B. Scott for their support. This work is supported by grants from the National Institutes of Health/National Cancer Institute through R01CA137060, R01CA139032, R01CA169458, R01CA172558 and R01CA157644 (to M.M.), ECOG-ACRIN grants CA180820 and CA180794 (to E.P.), grants from the Leukemia and Lymphoma Society (to M.M.), the California Institute for Regenerative Medicine through TR02-1816 (M.M.), and the William Lawrence and Blanche Hughes Foundation. M.M. is a Scholar of The Leukemia and Lymphoma Society and a Senior Investigator of the Wellcome Trust. M.R. is supported by the DFG through EXC294, TRR130 and SFB746.

Author Contributions Z.C. and M.M. designed experiments. M.M. conceived the study. Z.C., S.S., M.B., J.W.L., L.K. and E.P. performed experiments and interpreted data. H.G. performed statistical analysis. Y.X.T., A.S., J.U.J., J.E.C., S.B., T.W.M., A.L., H.J., M.R., A.W. and C.A.L. provided important reagents and mouse samples. E.P., S.P.H., C.L.W., A.M. and M.L.L. provided patient samples. B.T. and T.G.G. helped with gene expression analysis. The manuscript was written by Z.C. and M.M. and contributed to by all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.M. (markus.muschen@ucsf.edu).

METHODS

Patient samples and human cell lines. Patient samples (Supplementary Tables 1 and 2) were obtained with informed consent in compliance with Institutional Review Board regulations of the University of California San Francisco. Leukaemia cells from bone marrow biopsy of patients with Ph⁺ or Ph⁻ ALL were xenografted into sublethally irradiated NOD/SCID mice via tail vein injection. After passaging, leukaemia cells were harvested and cultured on top of OP9 stroma cells with minimum essential medium (MEM α ; Life Technologies) GlutaMAX without ribonucleotides and deoxyribonucleotides, supplemented with 20% FBS, 100 IU ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin and 1 mmol l⁻¹ sodium pyruvate. The human cell lines (Supplementary Table 3) were cultured in RPMI-1640 (Life Technologies) with GlutaMAX containing 20% FBS, 100 IU ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin at 37 °C in a humidified incubator with 5% CO₂. All of the human xenograft cells and cell lines are mycoplasma free.

Murine cell culture and BCR–ABL1 transformation. Bone marrow cells from constitutive or inducible knockout mice (for a list of genetic mouse models used in this study see Supplementary Table 4) were harvested and cultured in Iscove's modified Dulbecco's medium (IMDM; Invitrogen) with GlutaMAX containing 20% FBS, 50 μ M 2-mercaptoethanol, 100 IU ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin in the presence of cytokines. For pre-B-cell culture, bone marrow cells were cultured in IMDM with 10 ng ml⁻¹ recombinant mouse IL-7 (PeproTech) on OP9 stroma cells. For the ALL leukaemia model, pre-B cells were retrovirally transduced by BCR–ABL1. ALL cells generated from inducible knockout mice were retrovirally transduced with ER^{T2} or Cre-ER^{T2} virus, and puromycin selection was performed. 4-OHT was used to induce Cre-mediated gene deletion. For the CML-like leukaemia model, the myeloid-restricted protocol described previously was used²³, which generates CML-like cells. Briefly, bone marrow cells were cultured in IMDM with recombinant mouse IL-3 (10 ng ml⁻¹), IL-6 (25 ng ml⁻¹), SCF (50 ng ml⁻¹, PeproTech) and then transformed by BCR–ABL1 retrovirus. Cytokines were removed after BCR–ABL1 transduction.

In vivo transplantation of leukaemia cells. Murine pre-B ALL cells transformed by BCR–ABL1 were transduced with firefly luciferase retrovirus, selected with blasticidin, and then transduced with ER^{T2} or Cre-ER^{T2} virus, selected with puromycin. 4-OHT was used to induce Cre-mediated gene deletion for 24 h and 1 \times 10⁶ viable cells were injected into sublethally irradiated (250 cGy) NOD/SCID mice via the tail vein. For human leukaemia cells, a lentiviral vector encoding firefly luciferase was used. Bioimaging of leukaemia progression in mice was performed after transplantation with an *in vivo* IVIS 100 bioluminescence/optical imaging system (Xenogen). Fifteen minutes before measuring the luminescence signal, D-luciferin (Promega) prepared in PBS was injected intraperitoneally at an amount of 2.5 mg for each mouse. General anaesthesia was induced by using 5% isoflurane and continued during the process with 2% isoflurane given through a nose cone. When a mouse was terminally sick, it was euthanized and bone marrow and spleen cells were collected for flow-cytometry analysis. All mouse experiments were subject to institutional approval by the University of California San Francisco Institutional Animal Care and Use Committee. Six- to eight-week-old female NOD-SCID mice were randomly allocated into each treatment group. The minimal number of mice in each group was calculated by using the 'cpower' function in R/Hmisc package. No blinding was used.

Retroviral and lentiviral transduction. Retrovirus production was performed as described previously¹¹. Briefly, transfections of the retroviral constructs together with pHIT60 (*gag-pol*) and pHIT123 (*env*) were performed using Lipofectamine 2000 (Invitrogen). Sodium butyrate (10 mM) was used for induction. The virus supernatant was collected, filtered through a 0.45 μ m filter. For lentivirus, PCD/NL-BH (*gag-pol*) and pMN-VSV-G (*env*) were used for virus packaging. The lentivirus was concentrated by centrifugation filters from EMD Millipore. For transduction, non-tissue-culture-treated 6-well plates were coated with 50 μ g ml⁻¹ retronectin (Takara), and virus was loaded by centrifugation (2,000g, 90 min at 32 °C). Then virus was discarded and 2 \times 10⁶ pre-B cells were transduced per well by centrifugation at 600g for 30 min. Details of retroviral and lentiviral vectors used were provided in Supplementary Table 5.

Inhibitors. The BCR–ABL1 TKI imatinib was obtained from LC Laboratories. The INPP5D inhibitor 3AC and the Csk^{AS} inhibitor 3-IB-PP1 were obtained from EMD Millipore. The SYK TKI PRT062607 was purchased from Selleck Chemicals LLC.

Cell viability assay. One-hundred-thousand human ALL cells were seeded in a volume of 50 μ l medium in one well of a 96-well plate (BD Biosciences). Imatinib or any other inhibitor was diluted and incubated at the indicated concentration in a total volume of 100 μ l medium. After 3 days, cell counting kit-8 (Dojindo Molecular Technologies) was used to determine the number of viable cells. Fold changes were calculated using baseline values of vehicle treated cells as a reference (set to 100%).

Flow cytometry. Antibodies used in flow cytometry are mentioned in Supplementary Table 6. For cell cycle analysis, the BrdU flow cytometry kit (BD Biosciences) or Click-iT EdU Flow Cytometry Assay Kit (Invitrogen) was used according to the manufacturer's instructions. For evaluation of intracellular ROS levels, ALL cells were incubated for 7 min with 1 μ M 5-(and 6-)chloromethyl-2',7'-dichlorodihydrofluorescein diacetate (CM-H₂DCFDA; Invitrogen) at 37 °C for oxidation of the dye by ROS. After washing with PBS, the cells were incubated for an additional 15 min at 37 °C in PBS to allow complete deacetylation of the oxidized form of CM-H₂DCFDA by intracellular esterases. The levels of fluorescence were then directly analysed by flow cytometry, gated on viable cells.

Western blotting. CellLytic buffer (Sigma) supplemented with protease inhibitor cocktail (Roche Diagnostics) and phosphatase inhibitor cocktail set II (EMD Millipore) were used to lyse cells. Ten micrograms of protein lysates per sample were separated on mini precast gels (Bio-Rad) and transferred on nitrocellulose membranes (Bio-Rad). For the detection of proteins, primary antibodies, alkaline-phosphatase-conjugated secondary antibodies and chemiluminescent substrate (Invitrogen) were used. Details of primary antibodies are shown in Supplementary Table 7.

Colony-forming assay for mouse cells. Ten-thousand BCR–ABL1-transformed ALL cells or 100,000 CML-like cells were used for this assay. Cells were resuspended in murine MethoCult medium (StemCell Technologies) and plated on dishes (3 cm in diameter) with an extra dish of water to prevent evaporation. After 7 to 14 days, colonies were counted.

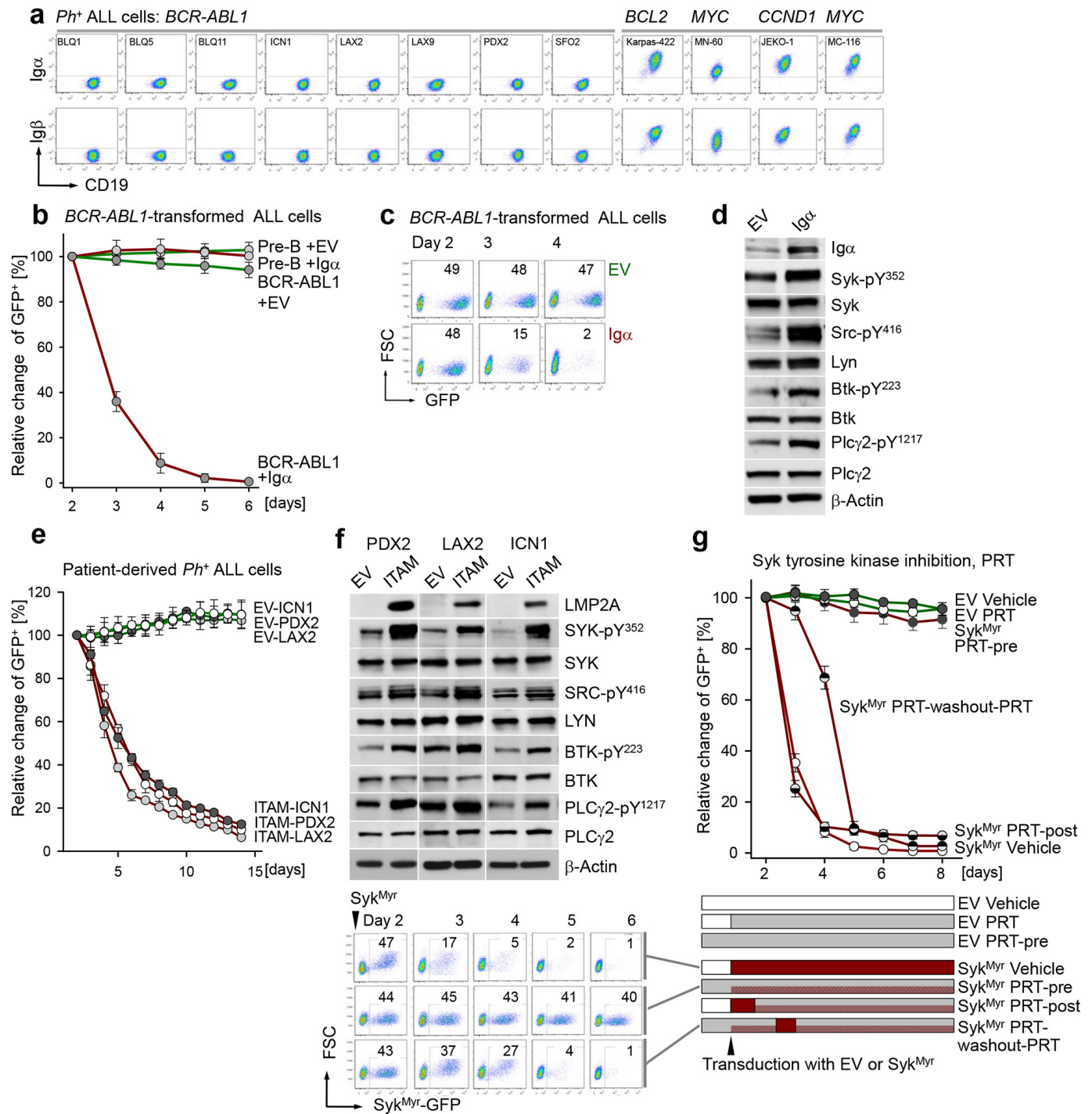
Senescence-associated β -galactosidase assay. This was performed on cytospin preparations as described previously¹¹.

DNA extraction and genotyping. Genomic DNA was extracted from mouse cells with NucleoSpin Tissue kit (MACHERY-NAGEL) and PCR was performed by using Taq DNA polymerase (NEB). The primer sequences are provided in Supplementary Table 8.

Gene expression and clinical outcome data. Clinical outcome and gene expression microarray data were derived from the National Cancer Institute TARGET Data Matrix (ftp://caftp.ncbi.nih.gov/pub/dcc_target/ALL/Phase_I/Discovery/clinical/) of the Children's Oncology Group (COG) Clinical Trial P9906 and from the ECOG Clinical Trial E2993. The end points of the clinical data include minimal residual disease (MRD) after 29 days of treatment (COG), overall survival (OS) and relapse-free survival (RFS) probability (COG and ECOG). Detailed information about the gene expression microarray data is provided in Supplementary Tables 9 and 10.

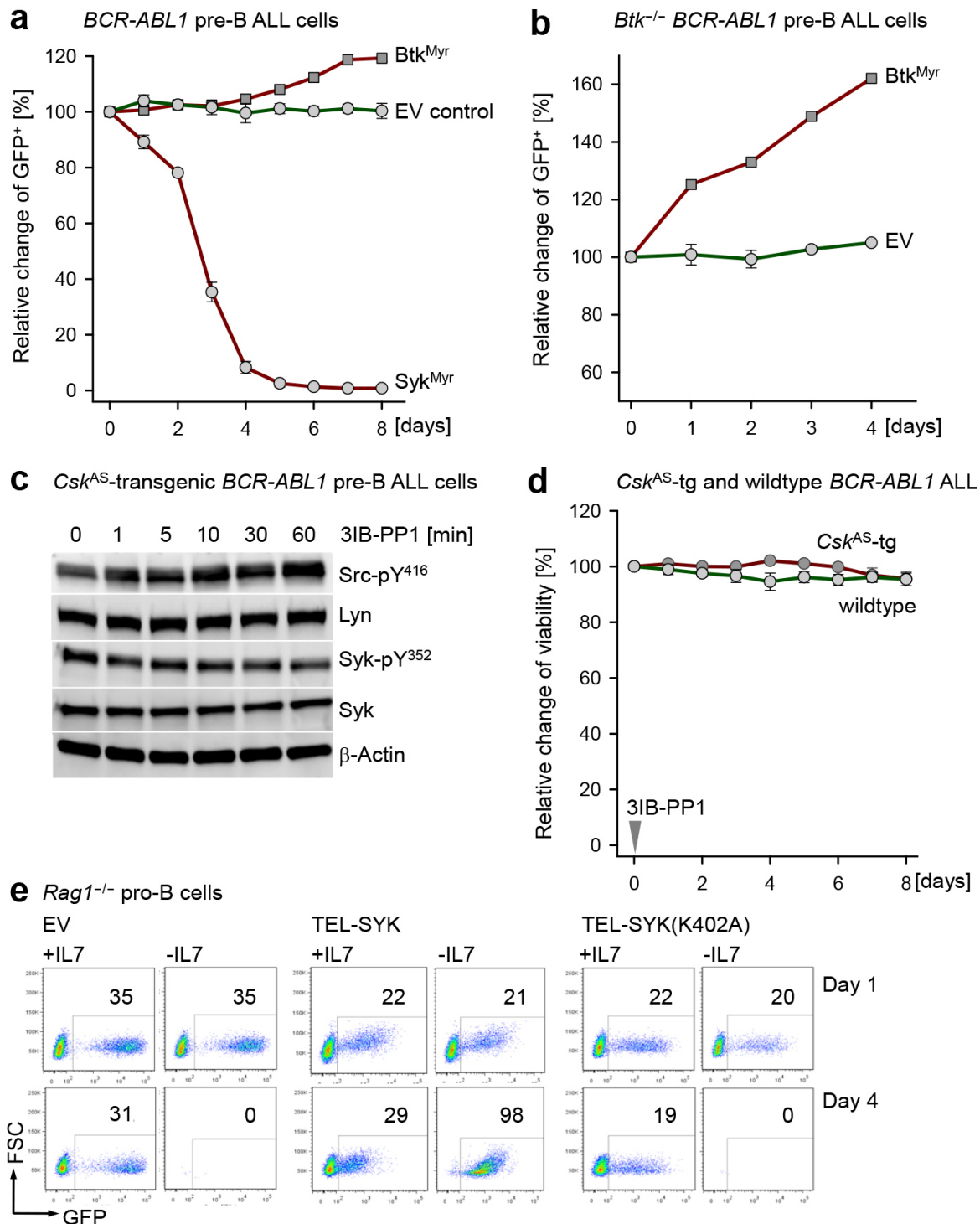
Statistical analysis. Unpaired, two-tailed Student's *t*-test was used to compare colony number, S-phase percentage and MFI of ITIM receptors between different groups. Two-sided Mann–Whitney Wilcoxon test was used to compare expression values between MRD⁺ versus MRD⁻ groups. OS or RFS probabilities were estimated using the Kaplan–Meier method. Log-rank test (two-sided) was used to compare patient survival between different groups. R package 'survival' version 2.35-8 was used for the survival analysis. In survival analysis, patients with ALL in each clinical trial (COG P9906 or ECOG E2993) were divided into two groups based on whether their expression was above or below the median level of a probeset or a gene (that is, the average of multiple probe sets for a gene). For a multiple-gene predictor (that is, a set of genes, such as in ITAM (*CD79A*, *CD79B*, *IHGM*) and ITIM (*PECAM1*, *LAIR1*, *CD300A*), the patients were split into four groups based on whether they had above or below the median expression levels of the sum of ITAM and the sum of ITIM gene expression levels: (1) ITAM^{High}ITIM^{Low} (\geq ITAM_{median} and $<$ ITIM_{median}), (2) ITAM^{High}ITIM^{High} (\geq ITAM_{median} and \geq ITIM_{median}), (3) ITAM^{Low}ITIM^{Low} ($<$ ITAM_{median} and $<$ ITIM_{median}), and (4) ITAM^{Low}ITIM^{High} ($<$ ITAM_{median} and \geq ITIM_{median}). Survival probabilities of the ITAM^{High}ITIM^{Low} versus ITAM^{Low}ITIM^{High} groups in the multiple-gene survival analysis were compared.

23. Li, S., Ilaria, R. L., Million, R. P., Daley, G. Q. & Van Etten, R. A. The P190, P210, and P230 forms of the BCR/ABL oncogene induce a similar chronic myeloid leukemia-like syndrome in mice but have different lymphoid leukemogenic activity. *J. Exp. Med.* **189**, 1399–1412 (1999).



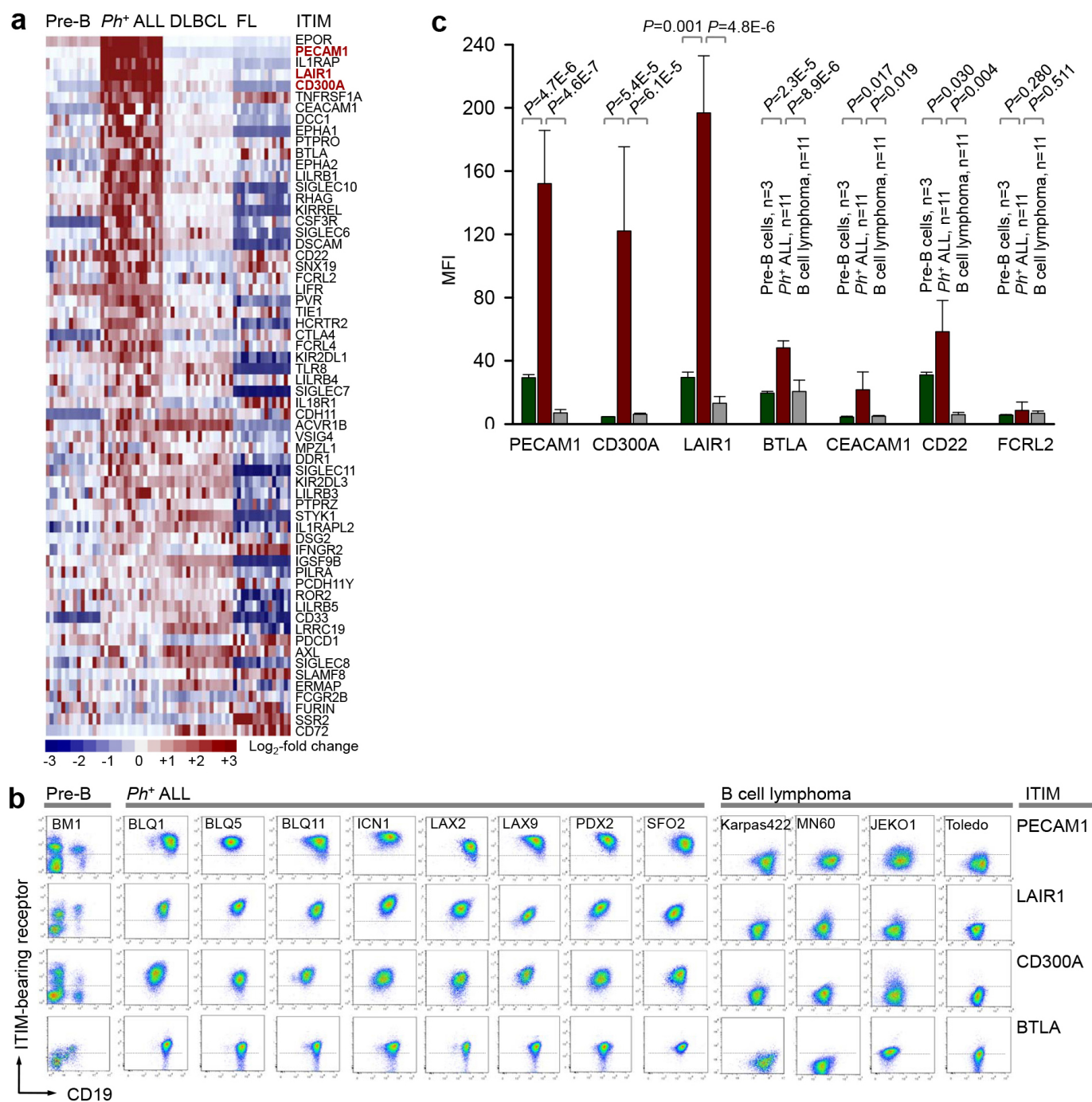
Extended Data Figure 1 | Reconstitution of ITAM signalling causes cell death in pre-B ALL. **a**, Flow cytometry staining for cell-surface Igα (CD79A) and Igβ (CD79B) was performed for patient-derived *Ph⁺* ALL cases ($n = 8$) and B-cell leukaemia/lymphomas lacking oncogenic tyrosine kinases ($n = 4$). **b**, **c**, Normal mouse pre-B cells or *BCR-ABL1*-transformed pre-B ALL cells were retrovirally transduced with CD8-Igα-GFP or empty vector (GFP) controls (EV). Relative changes of transduced (GFP⁺) populations were monitored by flow cytometry. **d**, Tyrosine phosphorylation of Syk, Src/Lyn, Btk and Plc-γ2 was studied in *BCR-ABL1* ALL cells that were transduced with Igα-GFP or GFP empty vector controls, using β-actin as loading control. Data (**c**, **d**) are representative of three independent experiments. **e**, Human *Ph⁺* ALL cells were transduced with GFP-tagged vectors for LMP2A-ITAM or empty

vector. Relative changes of transduced (GFP⁺) populations were monitored by flow cytometry ($n = 3$). **f**, LMP2A-ITAM or an empty vector was expressed in three cases of human *Ph⁺* ALL cells and effects on LMP2A expression and phosphorylation of SYK, SRC, BTK and PLC-γ2 were measured by western blot ($n = 3$). **g**, *BCR-ABL1*-transformed ALL cells were transduced with GFP-tagged Syk^{Myr} or an empty vector, and these cells were treated with the SYK inhibitor PRT ($2.5 \mu\text{mol l}^{-1}$) or vehicle either 1 day before transduction (PRT-pre), or 1 day after transduction (PRT-post), or pre-treated, then washed out for 1 day after transduction, and treated again with PRT. The relative changes of transduced (GFP⁺) cells were monitored by flow cytometry. Error bars represent means \pm s.d. from three independent experiments (**b**, **e**, **g**).



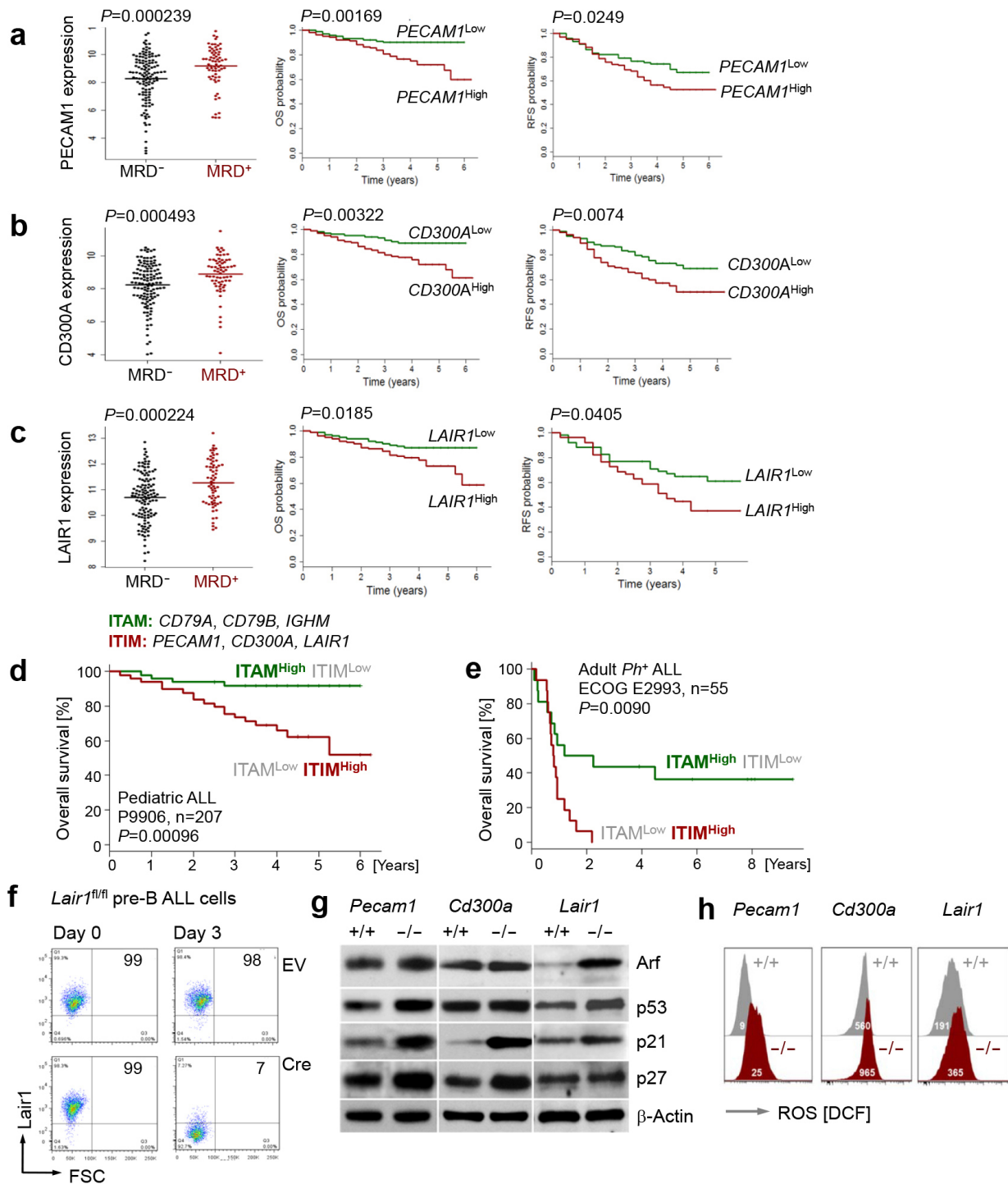
Extended Data Figure 2 | Reconstitution of proximal pre-BCR signalling in pre-B ALL cells. **a**, *BCR-ABL1*-transformed pre-B ALL cells were transduced with myristoylated (active) forms of Btk, Syk or empty vector controls (EV). Vectors were GFP-tagged and fractions of GFP⁺ cells were monitored over time. **b**, *Btk*^{-/-} *BCR-ABL1*-transformed pre-B ALL cells were transduced with myristoylated (active) Btk or empty vector (both tagged with GFP). Fractions of GFP⁺ cells were monitored over time. **c**, *Csk* is a negative regulator of Src family kinase activity. *Csk*-AS transgenic mice express an analogue (3IB-PP1) sensitive form instead of endogenous *Csk*. For inducible activation of Src kinase activity, we transformed pre-B cells from *Csk*-AS transgenic mice with *BCR-ABL1*. Addition of 3IB-PP1 (10 $\mu\text{mol l}^{-1}$) released *Csk*-mediated inhibition and induced increased phosphorylation of Src family kinases at Y416, but did not increase Syk Y352 phosphorylation (western blot). **d**, Pre-B cells from analogue-sensitive *Csk*-AS transgenic and wild-type mice were transformed

with *BCR-ABL1* and treated with 3IB-PP1. Cell viability in response to 3IB-PP1 treatment was monitored over time. Error bars (**a**, **b**, **d**) represent means \pm s.d. from three independent experiments. **e**, *Rag1*^{-/-} pro-B cells were expanded in the presence of 10 ng ml⁻¹ IL-7 and transduced with an empty vector (GFP), constitutively active SYK (TEL-SYK-GFP), or a kinase-dead mutant of SYK (TEL-SYK(K402A)-GFP). Then IL-7 was removed from cell cultures and the effect of IL-7 removal on cell viability was studied. *Rag1*^{-/-} pro-B cells transduced with empty vector or TEL-SYK(K402A)-GFP underwent apoptosis, whereas pro-B cells transduced with constitutively active SYK remained viable (data not shown). Four days after IL-7 removal, pro-B cells with constitutively active SYK had acquired growth factor (IL-7) independence, whereas pro-B cells with empty vector and TEL-SYK(K402A)-GFP remained dependent on IL-7. Data (**c**, **e**) are representative of three independent experiments.



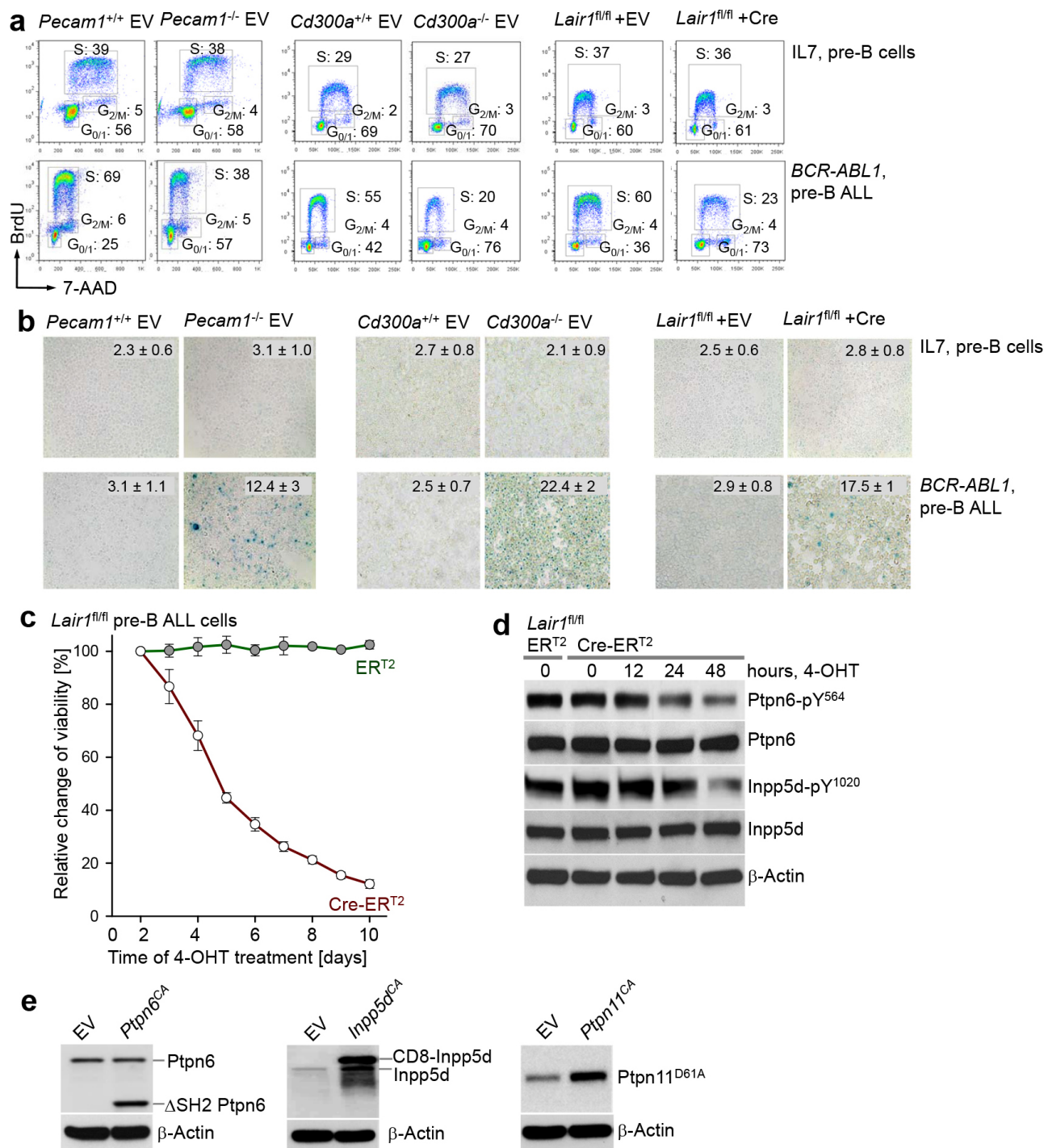
Extended Data Figure 3 | ITIM-bearing receptors are highly expressed on *Ph*⁺ ALL cells. **a**, Microarray data for 62 ITIM-bearing receptors are ranked based on the ratio of messenger RNA levels in *Ph*⁺ ALL compared to normal pre-B cells and mature B-cell lymphomas. **b**, Fluorescence-activated cell sorting (FACS) dot plots for double staining of PECAM1, CD300A, LAIR1 and BTLA with CD19 are shown for normal bone marrow pre-B cells ($n = 1$), *Ph*⁺ ALL cells ($n = 8$) and non-tyrosine-kinase-driven B-cell lymphoma ($n = 4$). **c**, Normal bone marrow mononucleated cells from bone marrow biopsies of healthy donors ($n = 3$), patient-derived *Ph*⁺ ALL ($n = 11$) and non-tyrosine-kinase-driven B-cell lymphoma ($n = 11$) were analysed by flow

cytometry for surface expression of the ITIM-bearing inhibitory receptors PECAM1, CD300A, LAIR1, BTLA, CEACAM1, CD22, FCRL2. Additional staining for CD72 and LILRB5 did not show significant differences between *Ph*⁺ ALL cells and normal pre-B cells (data not shown). Statistical analysis of mean fluorescence intensities (MFIs) for normal pre-B cells ($n = 3$), *Ph*⁺ ALL ($n = 11$) and non-tyrosine-kinase-driven B-cell lymphoma ($n = 11$) showed significantly increased expression levels of PECAM1, LAIR1 and CD300A in *Ph*⁺ ALL compared to normal pre-B cells and non-tyrosine-kinase-driven B-cell lymphoma. P values were calculated using unpaired, two-tailed Student's t -test.



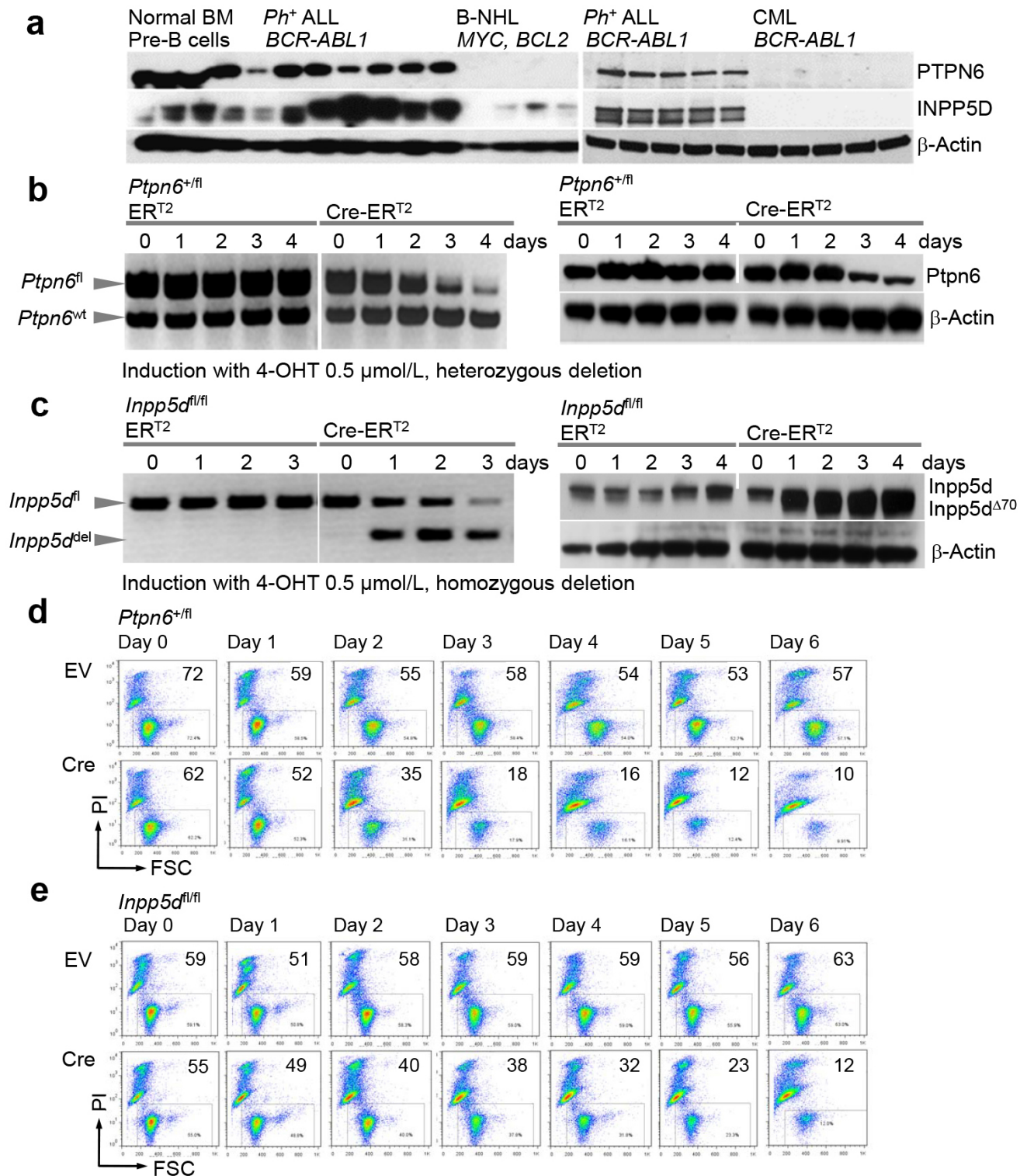
Extended Data Figure 4 | Higher than median expression levels of ITIM-bearing inhibitory receptors predict poor outcomes in patients with pre-B ALL. **a–c**, mRNA levels for PECAM1, CD300A and LAIR1 were measured in 207 patients with paediatric ALL (COG P9906). PECAM1, CD300A and LAIR1 mRNA levels for ALL cells from 124 patients that had no detectable minimal residual disease (MRD negative; black) on day 29 in their bone marrow were compared to mRNA levels in 67 patients with positive MRD (red) at the time of bone marrow biopsy (day 29). On the basis of higher or lower than median expression levels of PECAM1, CD300A and LAIR1, patients were segregated into two groups (High, $n = 104$; Low, $n = 103$; plots in middle and right). Overall survival (OS; middle) and relapse-free survival (RFS; right) probabilities were estimated by Kaplan–Meier survival analyses. P values were calculated by Mann–Whitney–Wilcoxon test (left panels; MRD status) and

log-rank test (middle and right panels; overall survival and relapse-free survival). **d, e**, ITAM-based agonists (CD79A, CD79B, IGHM) and ITIM-based inhibitors (PECAM1, CD300A, LAIR1) of pre-BCR signalling were combined into a six-gene outcome predictor based on ‘ITAM’ and ‘ITIM’ signatures and validated in two clinical trials for adults with Ph⁺ ALL (ECOG E2993) and children with ALL (COG P9906). P values were calculated by log-rank test. **f**, *Lair1* deletion was confirmed by flow cytometry. **g**, Expression of checkpoint molecules Arf, p53, p21 and p27 was measured by western blot in the presence and absence of *Pecam1* and *Cd300a* and upon inducible deletion of *Lair1* in BCR-ABL1 pre-B ALL cells. **h**, Accumulation of ROS was measured by staining with 2′,7′-dichlorofluorescein diacetate (DCF) in BCR-ABL1 pre-B ALL cells (grey histograms for control; red for gene deletion). Data are representative of three independent experiments (**f–h**).



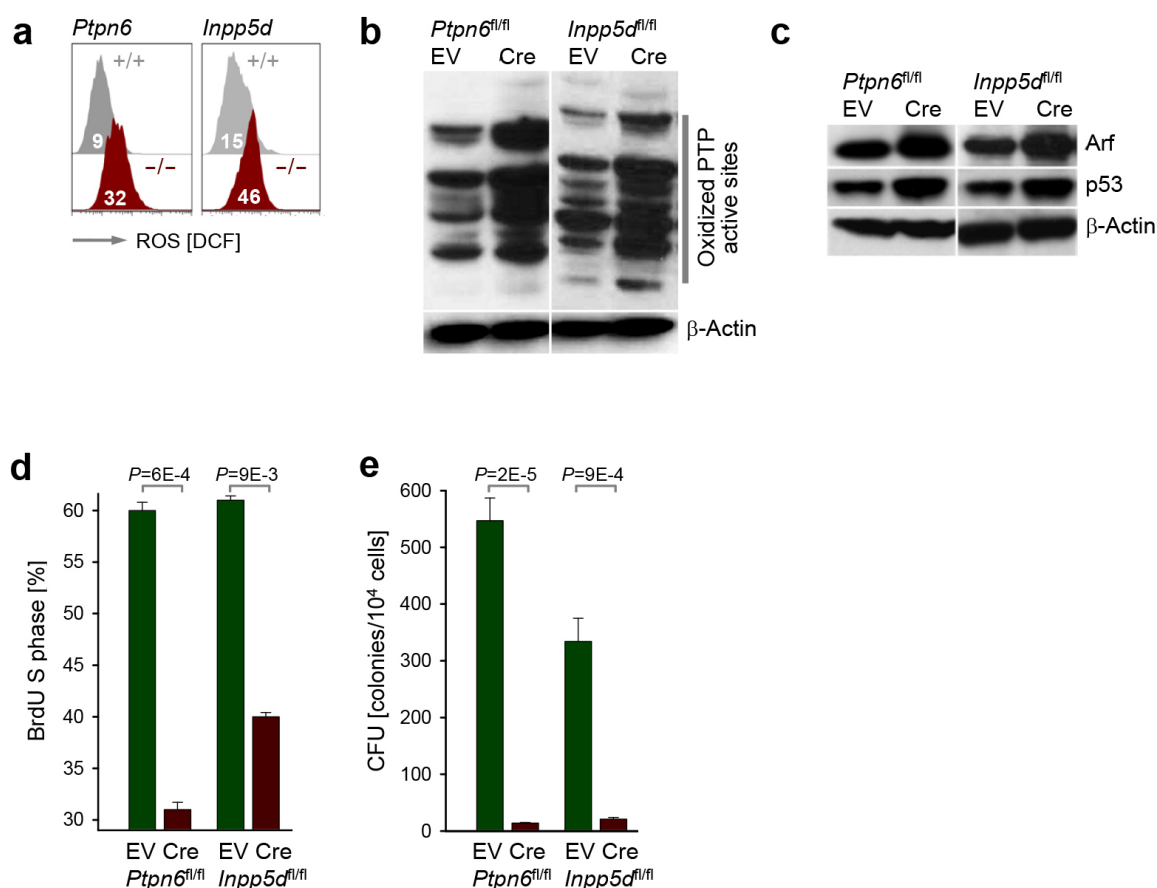
Extended Data Figure 5 | Consequences of genetic deletion of ITIM-bearing receptors in pre-B ALL cells. **a, b**, B-cell precursors from the bone marrow of *Pecam1*^{-/-} and *Cd300a*^{-/-} as well as *Lair1*^{fl/fl} mice and wild-type controls were propagated with IL-7 and transduced with an empty vector control (EV; normal B-cell precursors) or transformed with *BCR-ABL1* to model Ph⁺ ALL. *Lair1*^{fl/fl} pre-B and *BCR-ABL1* leukaemia cells were transduced with 4-OHT-inducible retroviral Cre. Cell cycle progression of normal pre-B cells (EV) and *BCR-ABL1* ALL cells was measured by BrdU staining (**a**). Propensity to cellular senescence was measured by staining of normal pre-B and *BCR-ABL1* ALL cells for senescence-associated β-galactosidase (**b**). **a**, Numbers indicate percentage of cells in each cell cycle phase. **b**, Numbers indicate percentage of β-galactosidase-positive cells. **c**, *Lair1*^{fl/fl} *BCR-ABL1* ALL cells were transduced with 4-OHT-inducible Cre (Cre-ER^{T2}) or an empty vector control (ER^{T2}).

Viability was measured by flow cytometry after 4-OHT treatment. **d**, Effects of inducible deletion of *Lair1* on phosphorylation levels of Ptpn6 and Inpp5d were measured by western blot. **e**, *Lair1*^{fl/fl} ALL cells were transduced with 4-OHT-inducible Cre. After antibiotic selection, ALL cells were transduced with a GFP-tagged empty vector control or GFP-tagged overexpression vectors for constitutively active forms of Ptpn6 (lacking autoinhibitory SH2 domain), Inpp5d (membrane-anchored by CD8) and Ptpn11 (constitutively active D61A mutation). Expression levels of Ptpn6, Inpp5d and Ptpn11 were measured by western blot using β-actin as loading control. The transduced cells were used for Cre-mediated deletion of *Lair1* to determine if expression of constitutively active Ptpn6, Inpp5d and Ptpn11 can rescue leukaemia cell survival. Data (**a, d, e**) are representative of three independent experiments. Data (**b, c**) represent means ± s.d. from three independent experiments.



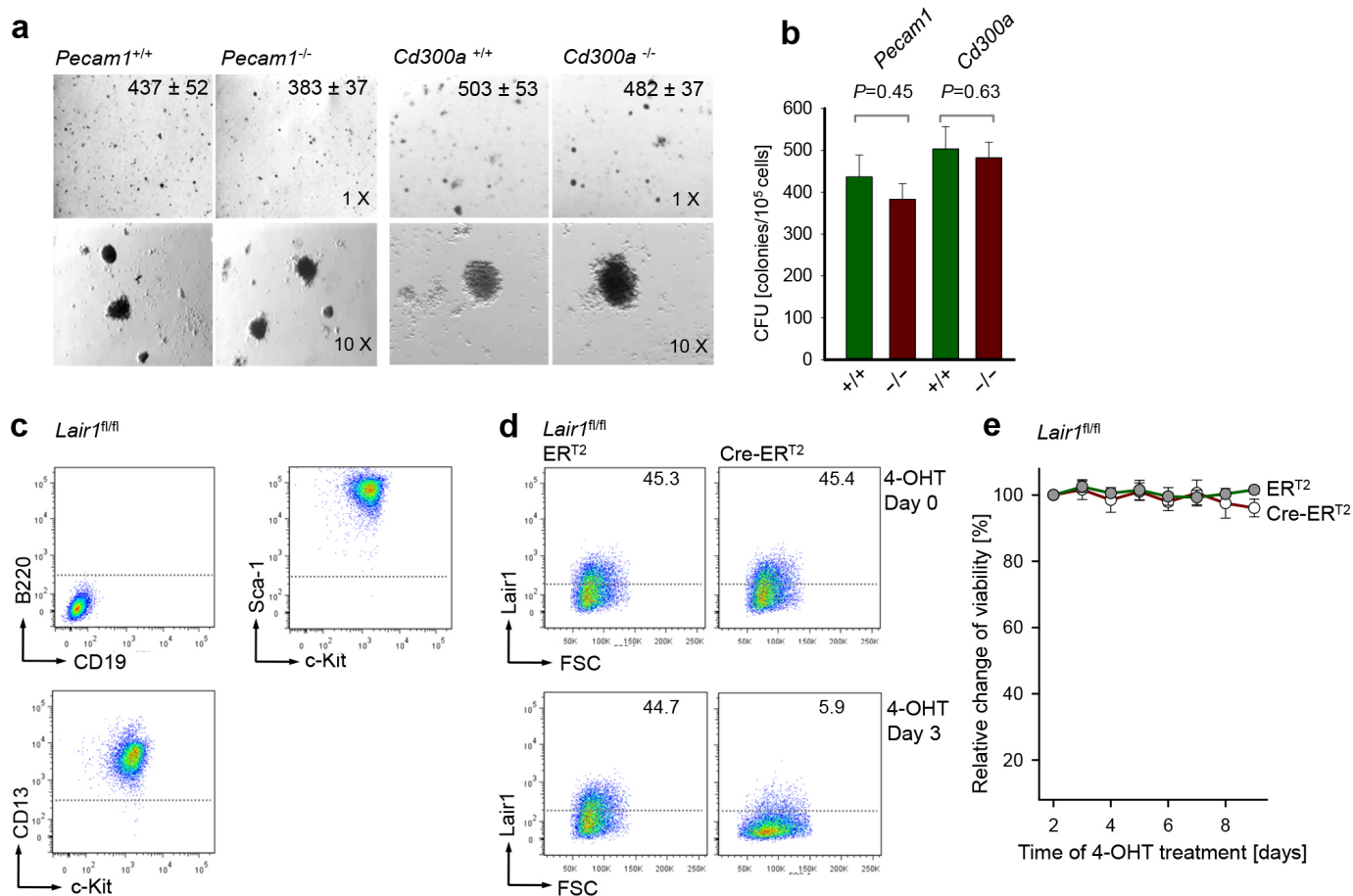
Extended Data Figure 6 | Inducible deletion of *Ptpn6* or *Inpp5d* causes cell death in pre-B ALL cells. **a**, Protein levels of PTPN6 and INPP5D were measured by western blot in CD19⁺ bone marrow pre-B cells from healthy donors ($n = 3$), patient-derived *Ph*⁺ ALL ($n = 8$) and B-cell leukaemia/lymphoma ($n = 4$) lacking an oncogenic tyrosine kinase. Additional western blot analyses compared expression levels of PTPN6 and INPP5D in patient-derived *Ph*⁺ ALL ($n = 5$) and patient-derived chronic phase CML cells ($n = 5$). **b**, **c**, Bone marrow cells were isolated from *Ptpn6*^{+/fl} or *Inpp5d*^{fl/fl} mice and pre-B cells were propagated with IL-7 (10 ng ml⁻¹). *Ptpn6*^{+/fl} and *Inpp5d*^{fl/fl} pre-B cells were then transduced with *BCR-ABL1* retrovirus and subsequently transduced with 4-OHT-inducible Cre (Cre-ER^{T2}) or an empty vector control (ER^{T2}). Addition of 4-OHT induced nuclear translocation of Cre and

Cre-mediated excision of *Ptpn6*^{+/fl} (one allele) and *Inpp5d*^{fl/fl} alleles as verified here by genomic PCR (**b**, left, for *Ptpn6*^{+/fl}; **c**, left, for *Inpp5d*^{fl/fl}). Near-complete deletion of the *Ptpn6*^{+/fl} (one allele) and *Inpp5d*^{fl/fl} floxed alleles was observed after 3 and 4 days, respectively, at the genomic level (left). Kinetics of protein depletion upon heterozygous deletion of *Ptpn6* and homozygous deletion of *Inpp5d* (*Inpp5d*^{Δ70}) was studied by western blot (right) using β-actin as loading control. **d**, **e**, Effects of Cre-mediated inducible deletion of *Ptpn6* (**d**) or *Inpp5d* (**e**) on *BCR-ABL1*-transformed pre-B ALL cell viability were measured by flow cytometry at the times indicated. Numbers denote percentages of viable cells (determined by forward scatter (FSC) and propidium iodide (PI) uptake). Data are representative of three independent experiments (**d**, **e**).



Extended Data Figure 7 | Functional consequences of inducible *Ptpn6* or *Inpp5d* deletion in pre-B ALL cells. **a**, The effects of deletion of *Ptpn6* or *Inpp5d* on cellular ROS levels were measured by flow cytometry using DCF in *BCR-ABL1* pre-B ALL cells (grey histograms for control; red for gene deletion). **b**, Whether ROS accumulation in response to deletion of *Ptpn6* or *Inpp5d* results in wide-spread cysteine-oxidation and, hence, inactivation, of multiple other PTP active sites was determined by western blot using antibodies against oxidized PTP active sites. **c**, Protein levels of the checkpoint molecules Arf and p53 were measured by western blot in *BCR-ABL1* ALL cells before (empty vector (EV)) and after (Cre) deletion of *Ptpn6* and *Inpp5d*. Data are

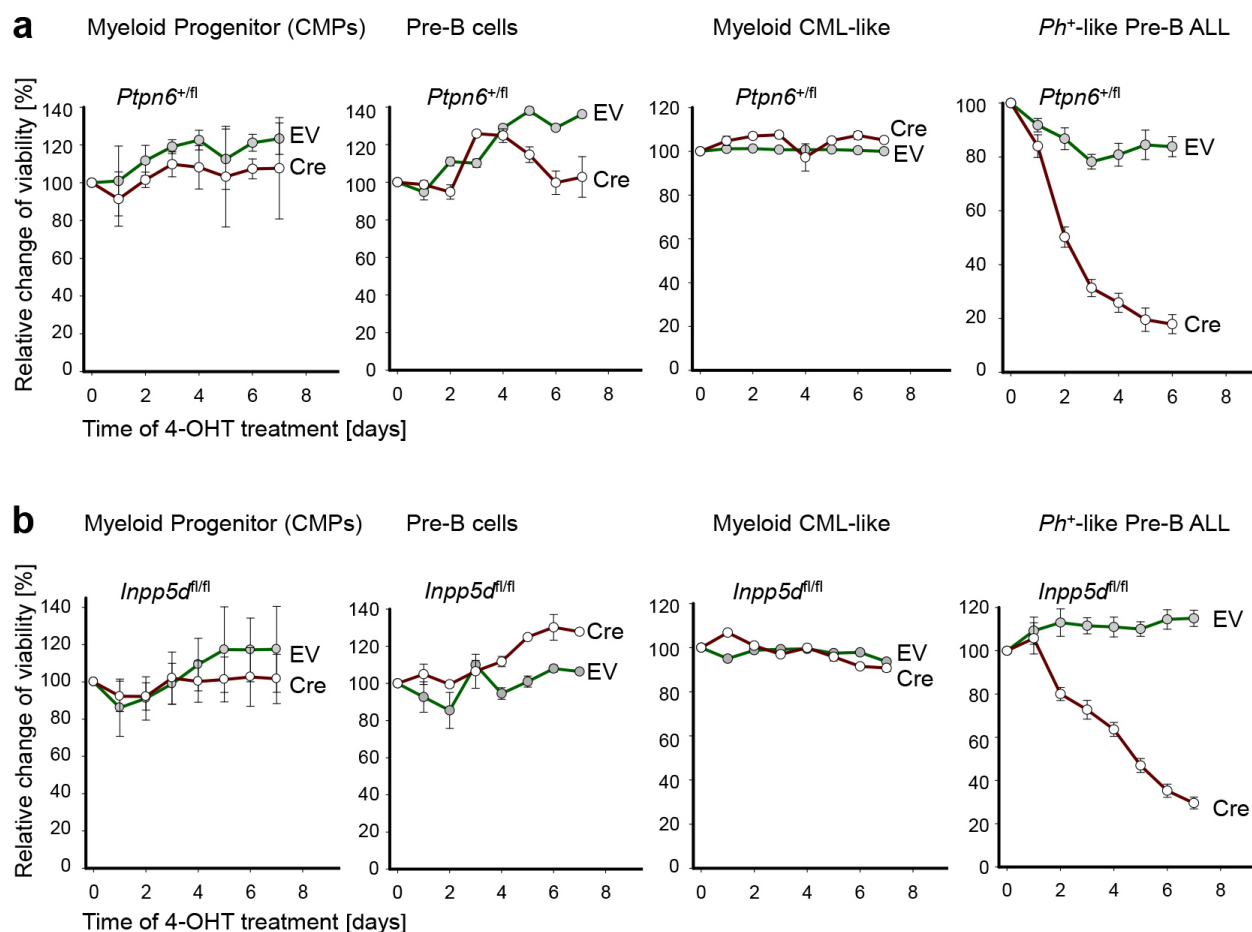
representative of three independent experiments (**a–c**). **d**, **e**, Functional readouts for inducible deletion of *Ptpn6* and *Inpp5d* include measurement of proliferation (BrdU incorporation) (**d**) and colony formation capacity in methylcellulose (colony-forming unit (c.f.u.) assay) (**e**). BrdU assays (flow cytometry) and c.f.u. data (images from colonies on plates) are presented in Fig. 3d, e. Quantitative and statistical analysis for BrdU incorporation (**d**) and c.f.u. assays (**e**) are depicted here as bar charts. *P* values were calculated by unpaired, two-tailed Student's *t*-test. Error bars (**d**, **e**) represent means \pm s.d. from three independent experiments.



Extended Data Figure 8 | Deletion of the ITIM-bearing receptors *Pecam1*, *Cd300a* or *Lair1* has no significant effects on myeloid CML-like cells.

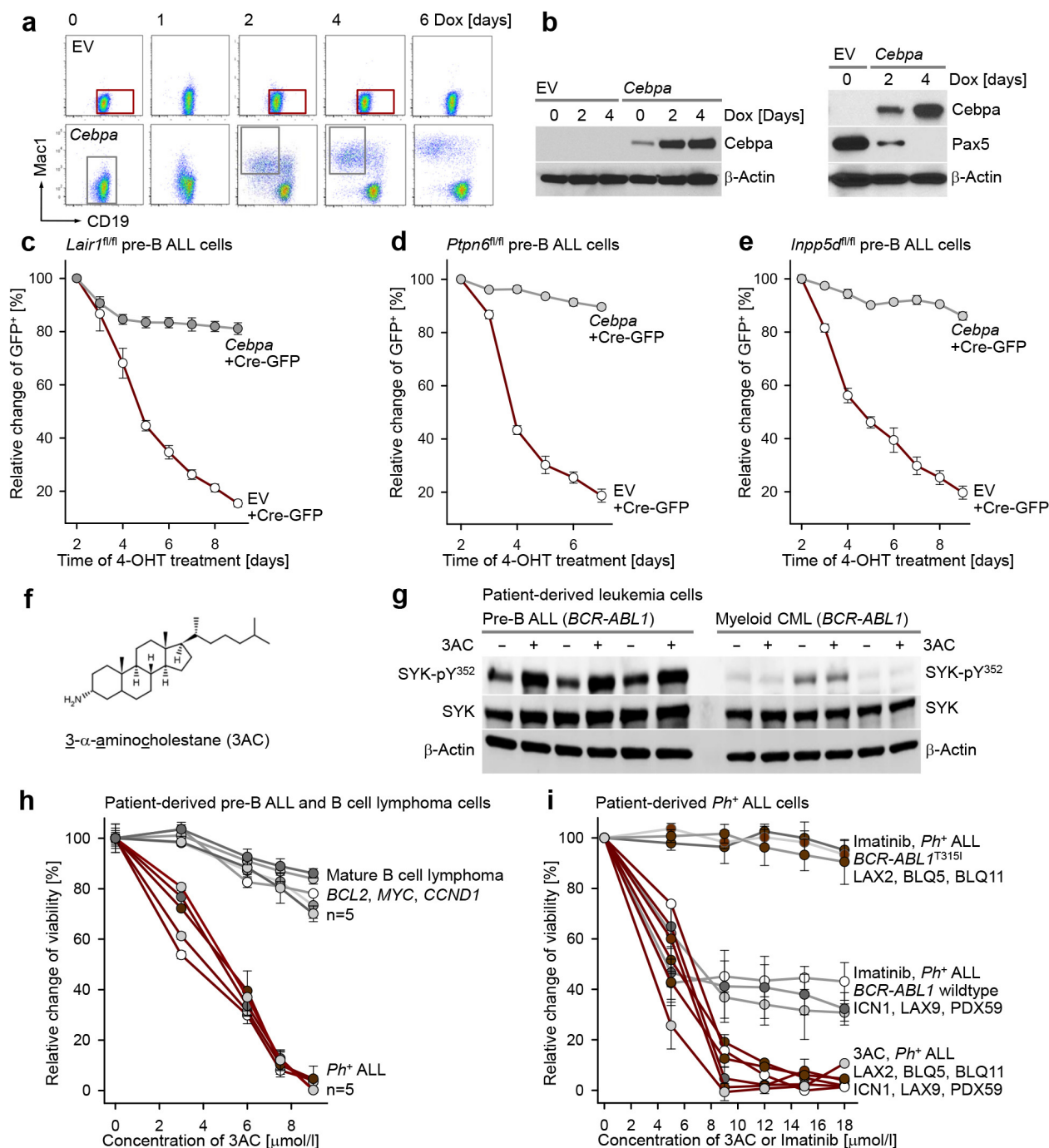
a, b, Myeloid progenitor cells from the bone marrow of *Pecam1*^{-/-} and *Cd300a*^{-/-} mice as well as age-matched wild-type controls were propagated in the presence of Il-3, Il-6 and Scf and transformed with retroviral *BCR-ABL1*. After 7 days, outgrowth of myeloid-lineage CML-like leukaemia was observed. One-hundred-thousand *Pecam1*^{-/-} and *Cd300a*^{-/-} CML-like cells as well as wild-type controls were plated in methylcellulose. Colonies were counted two weeks later (**a**). *P* values were calculated by unpaired, two-tailed Student's *t*-test (**b**). **c–e**, Myeloid progenitor cells from the bone marrow of *Lair1*^{fl/fl} mice

were propagated in the presence of Il-3, Il-6 and Scf and transformed with retroviral *BCR-ABL1*. After 7 days, outgrowth of myeloid-lineage CML-like leukaemia was observed and CML-like phenotype was verified by flow cytometry using antibodies against B220/CD19 (negative), Sca-1/c-Kit and CD13 (**c**). CML-like cells were transduced with 4-OHT-inducible Cre (Cre-ER^{T2}) and empty vector controls (ER^{T2}) and deletion of *Lair1* was verified by measurement of *Lair1* surface expression (**d**). After adding 4-OHT, cell viability of *Lair1*^{fl/fl} CML cells carrying ER^{T2} or Cre-ER^{T2} was monitored over 9 days by flow cytometry and is plotted in **e**. Data (**a**, **b**, **e**) represent means ± s.d. from three independent experiments.



Extended Data Figure 9 | Deletion of *Ptpn6* or *Inpp5d* specifically affects B-cell-lineage ALL cells but not normal pre-B cells, myeloid progenitors or myeloid leukaemia. a, b. Bone marrow mononuclear cells were isolated from $Ptpn6^{+/fl}$ and $Inpp5d^{fl/fl}$ mice. Myeloid progenitor cells were propagated with IL-6 (25 ng ml⁻¹), IL-3 (10 ng ml⁻¹) and Scf (50 ng ml⁻¹) and propagated as common myeloid progenitor cells (CMPs) or transformed with *BCR-ABL1* to induce myeloid CML-like leukaemia. Pre-B cells were expanded in the presence of IL-7 (10 ng ml⁻¹) and either propagated as pre-B-cell cultures or transformed by *BCR-ABL1* to induce Ph^+ ALL-like leukaemia. Lineage identity and

>95% purity of cell populations was verified by flow cytometry. $Ptpn6^{+/fl}$ and $Inpp5d^{fl/fl}$ CMPs, pre-B cells, CML-like and Ph^+ ALL-like leukaemia cells were then transduced with 4-OHT-inducible Cre (Cre) or an empty vector control (EV). Addition of 4-OHT induced nuclear translocation of Cre and Cre-mediated excision of *Ptpn6*^{+/fl} (one allele) (a) or *Inpp5d*^{fl/fl} alleles (b). Effects of inducible deletion on cell viability were measured by flow cytometry at the times indicated. Error bars (a, b) represent means \pm s.d. from three independent experiments.



Extended Data Figure 10 | The inhibitory receptor *Lair1* and the phosphatases *Ptpn6* and *Inpp5d* are specifically required by B-cell lineage leukaemia cells. **a, b**, B-cell lineage *BCR-ABL1* ALL cells were engineered with a doxycycline-inducible vector system for expression of *Cebpa*, which results in downregulation of B-cell antigens and myeloid-lineage differentiation as measured by flow cytometry (**a**) and western blot (**b**). Data (**a, b**) are representative of three independent experiments. **c–e**, *BCR-ABL1*-driven *Lair1^{fl/fl}*, *Ptpn6^{fl/fl}* and *Inpp5d^{fl/fl}* B-cell lineage ALL cells ($CD19^{+}$ $Mac1^{-}$) were reprogrammed into myeloid-lineage ($CD19^{-}$ $Mac1^{+}$) leukaemia cells by addition of doxycycline. Cell cultures were then transduced with 4-OHT-inducible GFP-tagged Cre and viability was measured in B-cell (gated on $CD19^{+}$ $Mac1^{-}$) and myeloid-lineage (gated on $CD19^{-}$ $Mac1^{+}$) populations. **f**, Structure of the INPP5D small-molecule inhibitor 3AC.

g, Patient-derived *Ph⁺* ALL ($n = 3$) and chronic-phase CML cells ($n = 3$) were treated with 3AC ($10 \mu\text{mol l}^{-1}$) for 15 min, and phosphorylation of SYK was measured by western blot, using β -actin as loading control. **h**, Dose-response curves are shown for five patient-derived cases of ALL (LAX2, LAX9, BLQ1, BLQ5 and PDX2, red curves) and five cases of B-cell leukaemia/lymphoma (lacking an oncogenic tyrosine kinase; KARPAS-422, MHH-PREB-1, JEKO-1, MN-60 and JJN-3, grey curves). **i**, Dose-response curves are shown for the treatment of six patient-derived cases of *Ph⁺* ALL that have acquired global resistance to TKI treatment (LAX2, BLQ5, BLQ11) or partial resistance (ICN1, LAX9, PDX59). Dose-response curves for the TKI imatinib are shown in grey and for the INPP5D inhibitor 3AC in red (concentration plotted on same scale for both agents). Error bars (**c–e, h–i**) represent means \pm s.d. from three independent experiments.

Lipid nanoparticle siRNA treatment of Ebola-virus-Makona-infected nonhuman primates

Emily P. Thi^{1*}, Chad E. Mire^{2,3*}, Amy C. H. Lee¹, Joan B. Geisbert^{2,3}, Joy Z. Zhou¹, Krystle N. Agans^{2,3}, Nicholas M. Snead¹, Daniel J. Deer^{2,3}, Trisha R. Barnard¹, Karla A. Fenton^{2,3}, Ian MacLachlan¹ & Thomas W. Geisbert^{2,3}

The current outbreak of Ebola virus in West Africa is unprecedented, causing more cases and fatalities than all previous outbreaks combined, and has yet to be controlled¹. Several post-exposure interventions have been employed under compassionate use to treat patients repatriated to Europe and the United States². However, the *in vivo* efficacy of these interventions against the new outbreak strain of Ebola virus is unknown. Here we show that lipid-nanoparticle-encapsulated short interfering RNAs (siRNAs) rapidly adapted to target the Makona outbreak strain of Ebola virus are able to protect 100% of rhesus monkeys against lethal challenge when treatment was initiated at 3 days after exposure while animals were viraemic and clinically ill. Although all infected animals showed evidence of advanced disease including abnormal haematology, blood chemistry and coagulopathy, siRNA-treated animals had milder clinical features and fully recovered, while the untreated control animals succumbed to the disease. These results represent the first, to our knowledge, successful demonstration of therapeutic anti-Ebola virus efficacy against the new outbreak strain in nonhuman primates and highlight the rapid development of lipid-nanoparticle-delivered siRNA as a countermeasure against this highly lethal human disease.

Historical Ebola virus (EBOV) outbreaks have previously ranged in size from a few to more than 400 cases, and were relatively well controlled by contact tracing and quarantine methods. In late 2013, an unprecedented outbreak caused by the Zaire species of EBOV began. This outbreak focused around the West African countries of Guinea, Liberia and Sierra Leone and has continued unabated for more than a year so far, with 25,213 cases and 10,460 deaths¹. Despite intensive containment efforts, the outbreak is still not under control and the need for medical countermeasures to both prevent and treat infections has never been greater.

While there are no approved vaccine or therapeutic treatment modalities available for preventing or managing EBOV infections, a few post-exposure approaches have demonstrated convincing efficacy against EBOV in a nonhuman primate (NHP) model that closely reproduces human infection. These include anti-EBOV monoclonal antibody administration alone (such as ZMapp) or with adenovirus-vectored interferon- α , and EBOV-targeting siRNAs encapsulated in lipid nanoparticles (LNPs) (TKM-Ebola) to potentiate cellular delivery^{3–5}. Several experimental treatments including ZMapp and TKM-Ebola have been employed under compassionate use protocols to treat small numbers of repatriated EBOV-infected medical staff in Europe and the United States². However, the contribution of these experimental treatments towards patient survival cannot be established, as several experimental treatments were applied in parallel alongside aggressive supportive care. Clinical trials have been initiated in West Africa to evaluate the efficacy of several experimental treatments including convalescent serum, vaccines, small molecules (brincidovir, now halted) and recently ZMapp, although these investigations may become hampered by the dwindling number of new cases of infection. Furthermore, up to now no treatments have been tested

against the current outbreak strain of EBOV under experimentally well-controlled conditions. Because much of the previous vaccine and antiviral development has been conducted in NHPs using the historical EBOV 1995 Kikwit strain from central Africa, there is a possibility that sequence changes documented in the West African strain^{6–8} may interfere with medical countermeasure efficacy, highlighting the need for treatments that can be rapidly adapted to mutated aetiological agents. While siRNA recognition is sequence dependent, adjustments for small viral nucleotide changes can be made rapidly. Monoclonal antibodies rely on cross-reactivity to conserved epitopes; if these are considerably changed, suitable antibodies must be identified *de novo*.

Sequence alignments of the nucleotide target sites of the TKM-Ebola siRNA cocktail, siEbola-2, with available sequences from the West African outbreak^{6–8} revealed conserved mismatches at antisense position 6 for siLpol-2 and at positions 3 and 15 for siVP35-2 that are not present in virus sequences endemic to central Africa (Fig. 1a). While certain positions within the prototypical siRNA structure are considered more crucial for function, and others better able to tolerate mismatches without erosion of activity, such effects are sequence-dependent and difficult to predict^{9–12}. Given this uncertainty, we took advantage of the rapid adjustment capability of the siRNA-LNP platform and designed a new siRNA cocktail, siEbola-3, in which these mismatches were corrected to enable full complementarity to West African outbreak EBOV sequences. We used a virus-free dual luciferase reporter assay to model the gene-silencing ability of the adjusted siRNA components against a representative central African strain versus the West African strain. Results demonstrated that the new siEbola-3 cocktail is fully active against the West African EBOV sequence, and retains activity against the central African sequence despite an impairment of the siVP35-3 siRNA component (Fig. 1b, see also Methods).

To assess medical countermeasure antiviral efficacy against the West African EBOV strain, we utilized *in vitro* and rhesus macaque models using a virus isolate from a lethal case in Guinea⁶. Deep sequencing of the challenge stock confirmed viral identity with 100% of the sequences containing the wild-type phenotype of 7 consecutive template uridines (7U) at the glycoprotein-editing site, confirming that viral virulence was not compromised during preparation of the challenge stock^{13,14}. It has been shown that macaques infected with 7U EBOV Kikwit succumb to infection earlier than those infected with 8U virus, and the protection afforded by some vaccine candidates decreases with EBOV 7U infection^{15,16}. Consistent with dual luciferase reporter predictions, both siEbola-2 and siEbola-3 LNPs were able to inhibit viral RNA levels in cultured cells infected with either EBOV Makona or EBOV Kikwit, although the siRNAs with full complementarity resulted in more activity (Extended Data Fig. 1).

siEbola-3 LNP treatment was able to protect NHPs against lethal challenge. NHPs were infected with the West African EBOV isolate and either left as untreated controls or administered siEbola-3 LNP beginning at 72 h after infection when animals were viraemic and

¹Tekmira Pharmaceuticals, Burnaby, British Columbia V5J 5J8, Canada. ²Galveston National Laboratory, University of Texas Medical Branch, Galveston, Texas 77550, USA. ³Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, Texas 77550, USA.

*These authors contributed equally to this work.

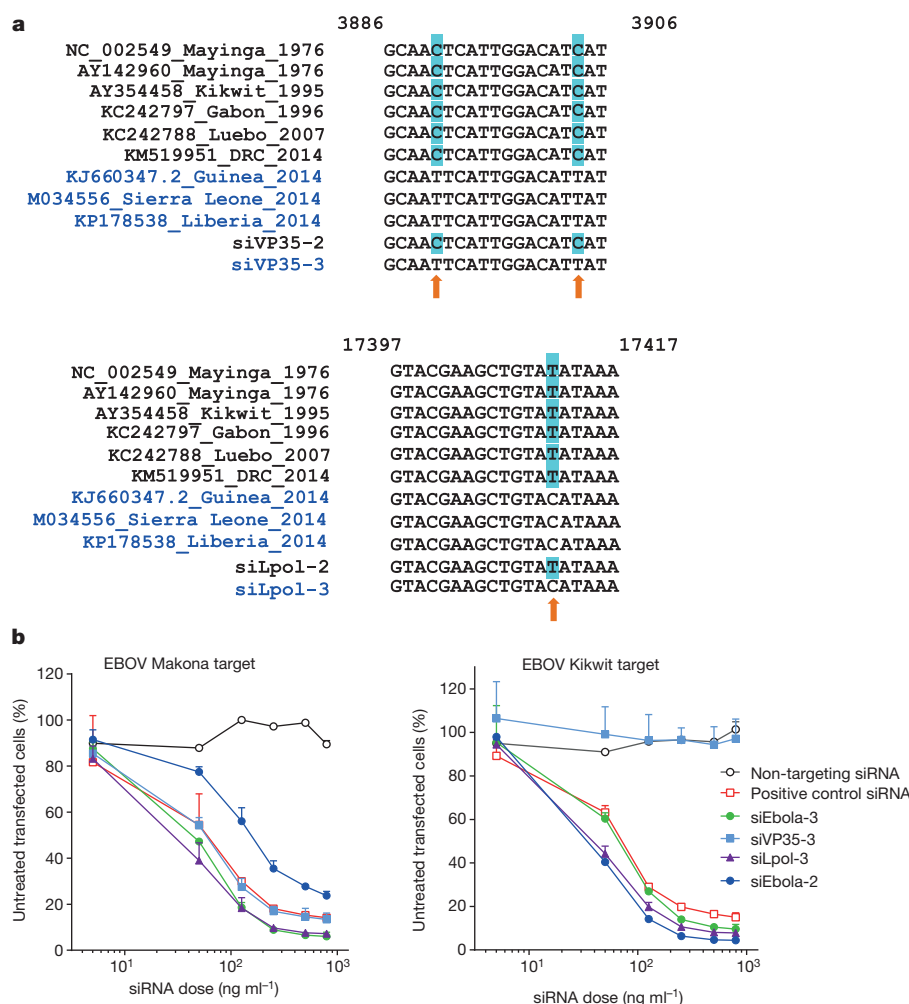


Figure 1 | siEbola-3 is active against EBOV Makona target sequences. **a**, The TKM-Ebola siRNA cocktail of siVP35-2 and siLpol-2 targets gene regions in central African EBOV sequences but West African outbreak sequences contain mutations at these locations. siEbola-3 has these target site mismatches corrected. **b**, siEbola-3 and its individual components, siVP35-3 and siLpol-3, are active against EBOV Makona sequences. Activity was assessed by dual luciferase reporter assay (see Methods). Shown is the *Renilla* luciferase/firefly luciferase ratio of each sample normalized to untreated cells. Results are mean \pm s.e.m. from one (negative control) or two (other data) biological replicates, conducted in technical triplicate.

clinically ill. All treated animals survived to study endpoint, while untreated control animals succumbed on days 8 and 9 (Fig. 2a). The time-to-death observed in untreated animals was similar to that reported after symptom onset in patients (9.8 ± 0.7 (mean \pm s.e.m.) days¹⁷), suggesting that EBOV infection in NHPs closely reproduces this aspect of human infection. Untreated control animals displayed mild clinical signs up until the day of euthanasia, after which a rapid deterioration of condition necessitated euthanasia (Fig. 2b). This is different from the disease course observed in NHPs infected with the EBOV Kikwit strain, in which animals tend to show a more gradual decline over the course of 1–3 days^{3,18,19} (Extended Data Table 1). In contrast to control animals, all siEbola-3-LNP-treated animals developed only transient mild clinical symptoms (Fig. 2b). Fever was observed in all infected animals with the exception of one treated animal, beginning at day 5 or 6 and continuing for 2–3 days until temperature returned to baseline (treated animals) or animals became hypothermic (Table 1). Petechial rashes were observed in all untreated and two treated animals, and these were milder than that seen previously in animals infected with EBOV Kikwit^{3,16,17}. Severe diarrhoea was also observed in two untreated animals infected with EBOV Makona, a clinical symptom associated with a fatal outcome in patients from this outbreak¹⁵. Diarrhoea is not as commonly observed in NHPs infected with EBOV Kikwit (T.W.G., unpublished observations). Taken together, these observations suggest that siEbola-3 LNP treatment protects against lethal EBOV Makona infection in a NHP model that recapitulates aspects of the disease observed in patients in the current West African outbreak, and that the disease manifestation in NHPs infected with EBOV Makona may differ from EBOV Kikwit infection.

siEbola-3 LNP treatment was also able to reduce viral load in infected animals (1–4 log-unit reductions in plasma viraemia when compared to

control animals, Fig. 2c), which correlated with 7.6- to 114-fold decreases in circulating viral genome detection (Fig. 2d, day 6). Peak viral RNA levels in untreated control animals were 8 and 9 log(viral copies per ml), respectively, well over the 10 million EBOV copies ml⁻¹ threshold associated with a higher fatality rate in patients¹⁷. At euthanasia, viral RNA was also widespread in tissues of untreated control animals, whereas it was only detected in the lymph nodes and spleen of treated animals at levels that were several magnitudes lower (Fig. 2e). These tissues were negative for infectious virus by plaque assay (data not shown), suggesting that the presence of viral RNA was not due to incomplete viral clearance. However, viral RNA detection at study endpoint in these sites of antigen presentation may reflect enforced viral replication in antigen presenting cells, which allows for adequate amounts of antigen to be presented to promote the adaptive immunity critical for survival after infection with a cytopathic virus²⁰. In accordance with this, immunohistochemical tissue evaluation showed positive EBOV antigen staining for the untreated control animals consistent with historical EBOV Kikwit-infected macaques^{3,18,19}, whereas detection of EBOV antigen in tissues of the fully recovered siEbola-3 LNP-treated animals was rare and limited to cells associated with antigen presentation (Fig. 3). No difference in viraemia levels was observed between EBOV Makona- and Kikwit-infected animals on the basis of limited available data (Extended Data Fig. 2a).

In conjunction with reductions in viral load, animals treated with siEbola-3 LNP showed moderate protection against liver dysregulation seen in untreated control animals infected with EBOV Makona, although the level of disturbance observed in infected animals was not as notable as those seen historically in rhesus macaques infected with EBOV Kikwit (Extended Data Fig. 2b–e). Treated animals also showed protection against EBOV-induced renal dysfunction as

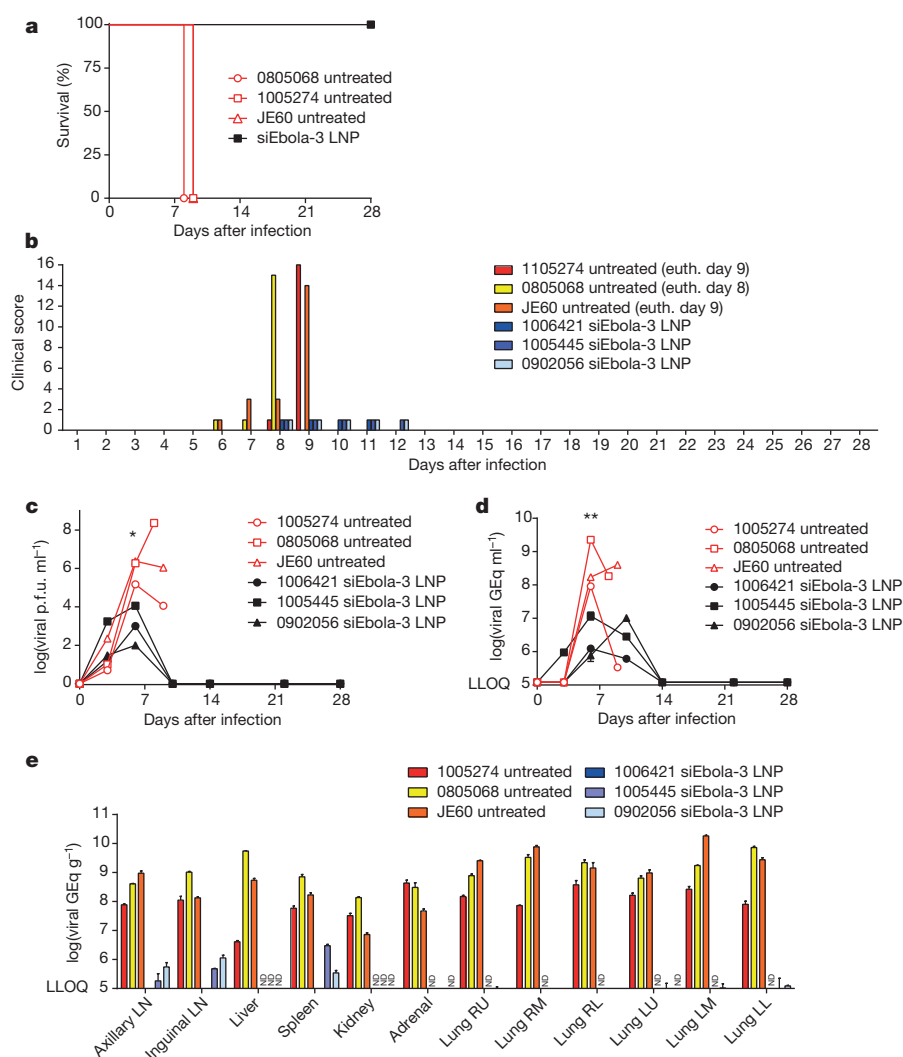


Figure 2 | siEbola-3 LNP treatment confers survival and reduces viral load. **a**, NHPs lethally challenged with EBOV Makona survive when treated with siEbola-3 LNP starting 72 h after infection. **b**, Clinical signs were improved in treated animals. Euth., euthanized. **c–e**, Treatment reduces infectious virus load (* $P = 0.0450$, one-sided t -test, day 6) (**c**), viral RNA in blood (** $P = 0.0023$, one-sided t -test, day 6) (**d**) and viral RNA in tissues (**e**). Lower limit of detection is 5 plaque-forming units (p.f.u.) per millilitre. **d, e**, Quantitative reverse transcriptase PCR (qRT-PCR) data shown are mean \pm s.d. of two technical replicates. GEq, genome equivalents; LLOQ, lower limit of quantitation ($4.8 \log_{10}(\text{viral copies per g})$ or $5.1 \log_{10}(\text{viral copies per ml})$); LL, left lower lobe; LM, left middle lobe; LN, lymph nodes; LU, left upper lobe; ND, not detected; RL, right lower lobe; RM, right middle lobe; RU, right upper lobe. $n = 3$ per group.

Table 1 | Clinical description and outcome of EBOV-challenged NHPs

Subject no.	Sex	Group	Clinical illness	Clinical pathology
0805068	M	Untreated control	Fever (d6); mild depression (d6–7); severe depression (d8); lethargy (d7–8); loss of appetite (d6–8); mild petechial rash (d8); rectorrhagia (d8); hunched posture (d6,7,8 morning.); recumbency (d8 pm); animal euthanized in afternoon on d8	Leukocytosis (d6,8); granulocytosis (d3,6,8); thrombocytopenia (d6,8); lymphopenia (d3,6,8); ALT >10-fold \uparrow (d8); AST >4-fold \uparrow (d6); AST >10-fold \uparrow (d8); ALP >2-fold \uparrow (d6); ALP >8-fold \uparrow (d8); GGT >10-fold \uparrow (d8); BUN >8-fold \uparrow (d8); CRE sevenfold \uparrow (d8); CRP >10-fold \uparrow (d6,8); fibrinogen >2-fold \uparrow (d6)
1105274	F	Untreated control	Fever (d6–7), mild depression (d8); severe depression (d9); lethargy (d8–9); loss of appetite (d8–9); mild petechial rash (d9); diarrhoea (d9); hunched posture (d9 am); recumbency (d9 pm); animal euthanized in afternoon on d9	Leukocytosis (d6); granulocytosis (d6); thrombocytopenia (d9); ALT >6-fold \uparrow (d9); AST >10-fold \uparrow (d9); BUN >2-fold \uparrow (d9); CRP >10-fold (d6); fourfold \uparrow (d9); APTT >2-fold \uparrow (d9); fibrinogen >2-fold \uparrow (d6)
JE60	M	Untreated control	Fever (d6), mild to moderate depression (d6–8); severe depression (d9); lethargy (d7–9); loss of appetite (d6–9); mild petechial rash (d6–9); severe epistaxis (d9); diarrhoea (d9); hunched posture (d6–8); recumbency (d9); animal euthanized in afternoon on d9	Thrombocytopenia (d6,9); lymphopenia (d6,9); hypoalbuminemia (d9); hypoproteinemia (d9); AST >6-fold \uparrow (d9); BUN >2-fold \uparrow (d9); CRP >10-fold (d6); 4-fold \uparrow (d9); APTT >2-fold \uparrow (d9)
0902056	F	72 h delay to treat	Fever (d8–10); mild depression (d8–12); loss of appetite (d5–13); mild petechial rash (d9–15); animal survived	Leukocytosis (d10); granulocytosis (d6,10); thrombocytopenia (d6,10,14); lymphopenia (d6); ALT >2-fold \uparrow (d6,10); AST >4-fold \uparrow (d6); AST >10-fold \uparrow (d10); GGT >2-fold \uparrow (d10); CRP >10-fold \uparrow (d6,10); fibrinogen >2-fold \uparrow (d6)
1005445	M	72 h delay to treat	Mild depression (d8–12); loss of appetite (d5–14); mild petechial rash (d9–13); animal survived	Granulocytosis (d10); Thrombocytopenia (d6,10); lymphopenia (d6); ALT >10-fold \uparrow (d6); ALT >5-fold \uparrow (d10); AST >10-fold \uparrow (d6,10); CRP >10-fold \uparrow (d6,10); APTT >3-fold \uparrow (10)
1006241	M	72 h delay to treat	Fever (d5–7); mild depression (d8–11); loss of appetite (d7–14); animal survived	Leukocytosis (d6,14); granulocytosis (d6,14); AST >7-fold \uparrow (d10); CRP >10-fold \uparrow (d6,10) fibrinogen >2-fold \uparrow (d6)

Days (d) after EBOV challenge are in parentheses. Fever is defined as a temperature more than 1.4°C over baseline or at least 0.8°C over baseline and $\geq 39.72^\circ\text{C}$. Mild rash: focal areas of petechiae covering less than 10% of the skin. Lymphopenia and thrombocytopenia are defined by a $\geq 40\%$ drop in numbers of lymphocytes and platelets, respectively. Leukocytosis and granulocytosis are defined by $\geq 40\%$ increase in numbers of white blood cells. ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; APTT, activated partial thromboplastin time; BUN, blood urea nitrogen; CRE, creatinine; CRP, C-reactive protein; F, female; GGT, gamma glutamyltransferase; M, male.

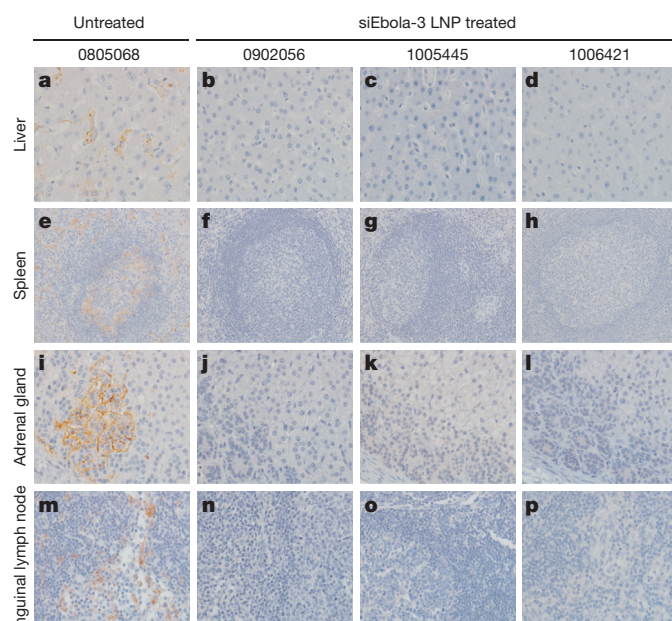


Figure 3 | EBOV Makona tissue pathology and antigen in NHPs untreated or treated with siEbola-3 LNP. **a–d**, Immunolabelling of sinusoidal lining and Kupffer cells in untreated animal. **b–d**, No immunolabelling in treated animals. **e**, Immunolabelling of dendriform mononuclear cells in red and white pulp of untreated animal. **f–h**, No immunolabelling in treated animals. **i**, Immunolabelling of cortical and interstitial cells in untreated animal. **j–l**, No immunolabelling in treated animals. **m**, Immunolabelling, dendriform mononuclear cells within subcapsular and medullary sinuses in untreated animal. **n–p**, No immunolabelling in treated animals. Original magnifications $\times 20$.

assessed by creatinine and blood urea nitrogen levels (Extended Data Fig. 2f, g). Smaller differences in coagulopathy, lymphopenia and thrombocytopenia were observed between treated and untreated animals (Table 1 and Extended Data Fig. 3). No differences in these parameters were apparent in untreated animals after infection with either EBOV Makona or EBOV Kikwit. Overall, these results indicate that siEbola-3 LNP treatment may confer additional protective benefits against clinical symptoms of EBOV-induced disease in addition to survival advantage and effective control of viral load. Some clinical pathology characteristics such as liver dysfunction were found to be not as profound in EBOV-Makona-infected NHPs when compared to that observed previously for EBOV Kikwit infection.

The current EBOV outbreak in West Africa highlights the need for antiviral therapeutics and prophylactics that can be readily and rapidly adapted to address the changing viral strain landscape. The use of a cocktail format (as opposed to a single siRNA) increases the likelihood of activity retention against newly emergent viral strains, as evidenced by the activity of siEbola-2 against EBOV Makona despite the presence of several nucleotide mismatches (Fig. 1b and Extended Data Fig. 1). Furthermore, the bipartite structure of TKM-Ebola, comprising both siRNA and LNP, allows for adjustments to the siRNA component to capitalize on emerging strain sequence data while maintaining the delivery functionality of the LNP component. Once viral sequence data are available, clinical grade drug product can be produced in as little as 8 weeks. Although TKM-Ebola (containing siEbola-2 and designed for central African EBOV) is currently under a US Food and Drug Administration (FDA) partial clinical hold regarding administration to healthy uninfected subjects, this product has been allowed by the FDA for use in cases of confirmed or suspected EBOV infection as the risk/benefit profile is quite different for patients facing a prospective high mortality rate compared to normal healthy individuals. The new siEbola-3 siRNA cocktail, shown here to possess robust activity against the latest EBOV Makona outbreak strain, is now being evaluated for efficacy in EBOV-infected patients in Sierra Leone, West Africa.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 January; accepted 1 April 2015.

Published online 22 April 2015.

- World Health Organization. *Ebola Data and Statistics*; <http://apps.who.int/gho/data/node.ebola-sitrep.quick-downloads?lang=en> (2015).
- Bishop, B. M. Potential and emerging treatment options for Ebola virus disease. *Ann. Pharmacother.* **49**, 196–206 (2015).
- Geisbert, T. W. *et al.* Postexposure protection of non-human primates against a lethal Ebola virus challenge with RNA interference: a proof-of-concept study. *Lancet* **375**, 1896–1905 (2010).
- Qiu, X. *et al.* Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature* **514**, 47–53 (2014).
- Qiu, X. *et al.* mAbs and Ad-vectored IFN- α therapy rescue Ebola-infected nonhuman primates when administered after the detection of viremia and symptoms. *Sci. Transl. Med.* **5**, 207ra143 (2013).
- Baize, S. *et al.* Emergence of Zaire Ebola virus disease in Guinea. *N. Engl. J. Med.* **371**, 1418–1425 (2014).
- Gire, S. K. *et al.* Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* **345**, 1369–1372 (2014).
- Kugelman, J. R. *et al.* Evaluation of the potential impact of ebola virus genomic drift on the efficacy of sequence-based candidate therapeutics. *mBio* **6**, e02227–14 (2015).
- Du, Q., Thonberg, H., Wang, J., Wahlestedt, C. & Liang, Z. A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res.* **33**, 1671–1677 (2005).
- Schwarz, D. S. *et al.* Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.* **2**, e140 (2006).
- Huang, H. *et al.* Profiling of mismatch discrimination in RNAi enabled rational design of allele-specific siRNAs. *Nucleic Acids Res.* **37**, 7560–7569 (2009).
- Dallas, A. *et al.* Inhibition of hepatitis C virus in chimeric mice by short synthetic hairpin RNAs: sequence analysis of surviving virus shows added selective pressure of combination therapy. *J. Virol.* **88**, 4647–4656 (2014).
- Volchkova, V. A., Dolnik, O., Martinez, M. J., Reynard, O. & Volchkov, V. E. Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. *J. Infect. Dis.* **204** (suppl. 3), S941–S946 (2011).
- Kugelman, J. R. *et al.* Ebola virus genome plasticity as a marker of its passaging history: a comparison of *in vitro* passaging to non-human primate infection. *PLoS ONE* **7**, e50316 (2012).
- Trefry, C. *et al.* *In vivo* pathological consequences of Ebola virus genome plasticity: is a 7U virus more lethal? *ASM Biodefense Emerging Dis. Res. Meeting*. Abstract 164 (2015).
- Hirschberg, R. *et al.* Challenges, progress, and opportunities: proceedings of the filovirus medical countermeasures workshop. *Viruses* **6**, 2673–2697 (2014).
- Schieffelin, J. S. *et al.* Clinical illness and outcomes in patients with Ebola in Sierra Leone. *N. Engl. J. Med.* **371**, 2092–2100 (2014).
- Geisbert, T. W. *et al.* Treatment of Ebola virus infection with a recombinant inhibitor of factor VIIa/tissue factor: a study in rhesus monkeys. *Lancet* **362**, 1953–1958 (2003).
- Hensley, L. E. *et al.* Recombinant human activated protein C for the postexposure treatment of Ebola hemorrhagic fever. *J. Infect. Dis.* **196** (suppl. 2), S390–S399 (2007).
- Honke, N. *et al.* Enforced viral replication activates adaptive immunity and is essential for the control of a cytopathic virus. *Nature Immunol.* **13**, 51–57 (2011).

Acknowledgements We thank V. Borisevich for assistance with clinical pathology assays performed in the GNL BSL-4 laboratory. We also thank S. Klassen for his assistance with siRNA-LNP preparation. This study was supported by the Department of Health and Human Services, National Institutes of Health grant U19AI109711 to T.W.G. and I.M.

Author Contributions E.P.T. and J.Z.Z. designed the siRNA and did preparative work for dual luciferase reporter studies. E.P.T., N.M.S., J.Z.Z. and A.C.H.L. designed the dual luciferase reporter studies. J.Z.Z., T.R.B. and N.M.S. conducted the dual luciferase reporter studies and analysed the data. E.P.T., C.E.M., A.C.H.L., I.M. and T.W.G. conceived and designed the NHP study. C.E.M., J.B.G., D.J.D. and T.W.G. performed the NHP challenge and treatment experiments and conducted clinical observations of the animals. J.B.G., K.N.A. and D.J.D. performed the clinical pathology assays. J.B.G. performed the EBOV Makona infectivity assays. C.E.M. and K.N.A. performed the PCR assays. E.P.T., C.E.M., J.B.G., K.N.A., D.J.D., K.A.F., A.C.H.L., I.M. and T.W.G. analysed the data. K.A.F. performed histological and immunohistochemical analysis of the data. E.P.T., C.E.M., A.C.H.L. and T.W.G. wrote the paper. All authors had access to all of the data and approved the final version of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of the paper. The other authors declare competing interests: details are available in the online version of the paper. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the University of Texas Medical Branch. Correspondence and requests for materials should be addressed to T.W.G. (twgeisbe@utmb.edu).

METHODS

Dual luciferase reporter assay. The psiCHECK2 (Promega) vector was used to construct the EBOV Makona and EBOV Kikwit strain reporter plasmids used in this study (Genscript). In brief, to construct the EBOV Makona or EBOV Kikwit reporter plasmids, two 201-base-pair (bp) regions of either the EBOV Makona strain or EBOV Kikwit strain genomes containing the VP35 and Lpol target sites (nucleotide positions 17287–17488, and 3817–4018 of GenBank accession number KJ660347.2 or AY354458) were fused together and cloned into the 3' untranslated region (UTR) of the *Renilla* luciferase gene between the XhoI and NotI restriction sites to allow for the detection of siRNA activity as represented by decreased *Renilla* luciferase activity. siLpol-3 and siVP35-3 were synthesized at ST Pharm, and siLpol-2 and siVP35-2 were synthesized at Integrated DNA Technologies. Individual duplexes and the siEbola-3 or siEbola-2 cocktail (1:1 molar mixture of siLpol-3 and siVP35-3 or siLpol-2 and siVP35-2, respectively) were encapsulated in LNP by the process of spontaneous vesicle formation as previously reported²¹. The resulting LNPs were dialysed against PBS and sterilized through a 0.2-µm filter before use. siRNAs targeting *Renilla* luciferase and MARV NP (synthesized by Integrated DNA Technologies) were also encapsulated in LNP and were included as positive and negative controls, respectively.

Authenticated HepG2 cells were obtained from ATCC (ATCC HB-8065). Cells were tested for mycoplasma before experimentation. HepG2 cells were transfected with the EBOV Makona or EBOV Kikwit psiCHECK2 plasmid construct using Lipofectamine 2000 (Life Technologies) and treated with siRNA-LNP at 5, 50, 125, 250, 500 and 750 ng ml⁻¹. Transfected cells were incubated for 24 h, followed by measurement of *Renilla* and firefly luciferase activities using a luminometer. Results were expressed as a percentage of the *Renilla*/firefly luciferase activity in cells transfected with the reporter plasmid only (no siRNA treatment).

EBOV Makona virus and sequence analysis. The EBOV Makona strain seed stock originated from serum from a fatal case during the 2014 outbreak in Guékédou, Guinea (Zaire ebolavirus isolate *Homo sapiens*-wt/GIN/2014/Makona-Guekedou-C07, accession number KJ660347.2)⁶ and was passaged twice in authenticated Vero E6 cells obtained from ATCC (ATCC, CRL-1586). The cells were not tested for mycoplasma. The EBOV Makona strain passage 2 seed stock was extracted in Trizol LS (Invitrogen) then purified using Zymo Research Direct-zol RNA mini-prep (Zymo Research) per manufacturer's instructions. Complementary DNA was generated from purified RNA using the Ovation RNA-seq 2 kit, which was subsequently used for the preparation of the double-stranded DNA library using the Encore Ion Torrent library prep kit (NuGen). Sequencing was performed by the UTMB Molecular Core on the Ion Torrent using 318-v2 deep sequencing chips. Sequence analysis was performed using Seqman NGen software (DNA Star) based on paired-end analysis of 100-bp overlaps.

In vitro infections. HepG2 cells (ATCC HB-8065) were seeded at 1E05 cells/well in 24-well culture plates and incubated at 37 °C/5% CO₂ overnight before infection with 0.1 multiplicity of infection (MOI) of either EBOV Makona or Kikwit. Cells were incubated with virus for 1 h, then washed four times with PBS and treated with siRNA-LNP at 51.2, 6.4 and 0.8 ng ml⁻¹. Cells were incubated for 48 h after treatment before collecting cell supernatants for RNA extraction by Trizol and qRT-PCR assessment.

Animal challenge. Six healthy adult rhesus macaques (*Macaca mulatta*) of Chinese origin (4–8 kg, three males and three females, 4–8 years old) were inoculated intramuscularly with 1,000 p.f.u. of EBOV Makona strain. The historical EBOV Kikwit data was obtained from six healthy rhesus macaques (six females, 4–8 years old) inoculated intramuscularly with 1,000 p.f.u. of EBOV Kikwit strain. Sample sizes were based on the availability of rhesus macaques. Animals were randomized with Microsoft Excel into treatment or control groups. siEbola-3 LNP (0.5 mg kg⁻¹) was administered to three of the EBOV-Makona-infected macaques by bolus intravenous infusion 72 h after EBOV challenge while the control animals were not treated. The three treated animals received additional treatments of siEbola-3 LNP on days 4, 5, 6, 7, 8 and 9 after EBOV challenge. All animals (six infected with EBOV Makona and six infected with EBOV Kikwit) were given physical examinations and blood was collected at the time of challenge and on days 3, 6, 10, 14, 22 and 28 after EBOV challenge or at time of euthanasia. In addition, all animals were monitored daily and scored for disease progression with an internal filovirus scoring protocol approved by the UTMB Institutional Animal

Care and Use Committee. The scoring changes measured from baseline included posture/activity level, attitude/behaviour, food and water intake, weight, respiration and disease manifestations such as visible rash, haemorrhage, ecchymosis or flushed skin. A score of ≥9 indicated that an animal met criteria for euthanasia. This study was not blinded.

Detection of viraemia and viral RNA. RNA was isolated from whole blood or tissues using the Viral RNA Mini Kit or RNeasy Kit (Qiagen) using 100 µl of blood into 600 µl of buffer AVL, or 100 mg of tissue per manufacturer's instructions, respectively. Primers/probe targeting the VP30 gene of EBOV were used for qRT-PCR with the probe used here being 6-carboxyfluorescein (6FAM)-5'-CCGT CAATCAAGGAGCGCCTC3'-6 carboxytetramethylrhodamine (TAMRA) for the EBOV Makona NHP and EBOV Makona and Kikwit *in vitro* studies (Life Technologies). EBOV RNA was detected using the CFX96 detection system (BioRad Laboratories) in One-step probe qRT-PCR kits (Qiagen) with the following cycle conditions: 50 °C for 10 min, 95 °C for 10 s, and 40 cycles of 95 °C for 10 s and 59 °C for 30 s. Threshold cycle (C_t) values representing EBOV genomes were analysed with CFX Manager Software, and data are shown as mean ± s.d. of technical replicates. To create the genome equivalent standard, RNA from EBOV stocks was extracted and the number of EBOV genomes calculated using Avogadro's number and the molecular mass of the EBOV genome.

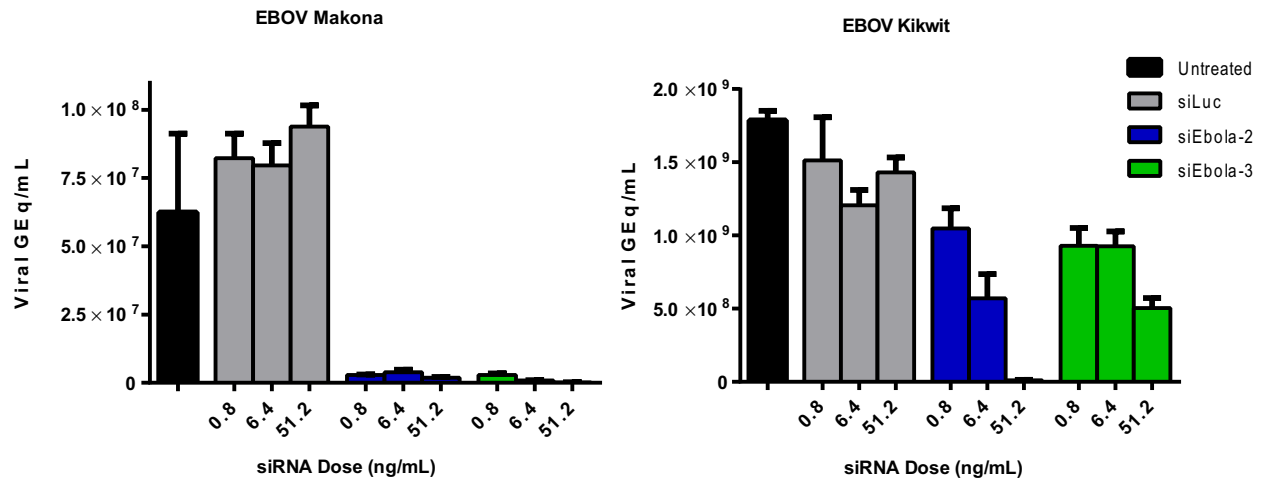
Virus titration was performed by plaque assay with Vero E6 cells from all serum samples as previously described^{3,16,17}. In brief, increasing tenfold dilutions of the samples were adsorbed to Vero E6 monolayers in duplicate wells (200 µl); the limit of detection was 5 p.f.u. ml⁻¹.

Haematology, serum biochemistry and blood coagulation. Total white blood cell counts, white blood cell differentials, red blood cell counts, platelet counts, haematocrit values, total haemoglobin concentrations, mean cell volumes, mean corpuscular volumes and mean corpuscular hemoglobin concentrations were analysed from blood collected in tubes containing EDTA using a laser-based haematological analyser (Beckman Coulter). Serum samples were tested for concentrations of albumin, amylase, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma-glutamyltransferase, glucose, cholesterol, total protein, total bilirubin, blood urea nitrogen, creatinine and C-reactive protein by using a Piccolo point-of-care analyser and Biochemistry Panel Plus analyser discs (Abaxis). Citrated plasma samples were analysed for coagulation parameters prothrombin time, activated partial thromboplastin time, and fibrinogen on the Start4 instrument using the PTT Automate, STA Neoplastine CI plus, and Fibrin-Prest Automate, kits, respectively (Diagnostica Stago).

Histopathology and immunohistochemistry. Necropsy was performed on all subjects. Tissue samples of all major organs were collected for histopathological and immunohistochemical examination, immersion-fixed in 10% neutral buffered formalin, and processed for histopathology as previously described^{16,17}. For immunohistochemistry, specific anti-EBOV immunoreactivity was detected using an anti-EBOV VP40 protein rabbit primary antibody (Integrated BioTherapeutics) at a 1:4,000 dilution. In brief, tissue sections were processed for immunohistochemistry using the Dako Autostainer (Dako). Secondary antibody used was biotinylated goat anti-rabbit IgG (Vector Laboratories) at 1:200 followed by Dako LSAB2 streptavidin-HRP (Dako). Slides were developed with Dako DAB chromagen (Dako) and counterstained with haematoxylin. Non-immune rabbit IgG was used as a negative control. Liver, adrenal gland and inguinal lymph nodes representative images were taken at ×40 magnification, and spleen taken at ×20 magnification from control animal 0805068 (Fig. 3a, e, h, m) or treated animals 0902056 (Fig. 3b, f, j, n), 1005445 (Fig. 3c, g, k, o), and 1006421 (Fig. 3d, h, l, p).

Statistical analyses. Analysis was conducted with Graphpad Prism software (version 6.04). A paired *t*-test (one-sided) was used to compare untreated and treated group means on day 6 for qRT-PCR (untreated group mean ± s.d. was 8.51 log(GE q ml⁻¹) ± 0.74; siEbola-3 LNP treated group was 6.36 log(GE q ml⁻¹) ± 0.62) and viraemia (untreated group mean ± s.d. was 5.94 log(p.f.u. ml⁻¹) ± 0.67; siEbola-3 LNP-treated group was 3.02 log(p.f.u. ml⁻¹) ± 1.03). No statistical methods were used to predetermine sample size.

21. Ma, H. *et al.* Formulated minimal-length synthetic small hairpin RNAs are potent inhibitors of hepatitis C virus in mice with humanized livers. *Gastroenterology* **146**, 63–66 (2014).

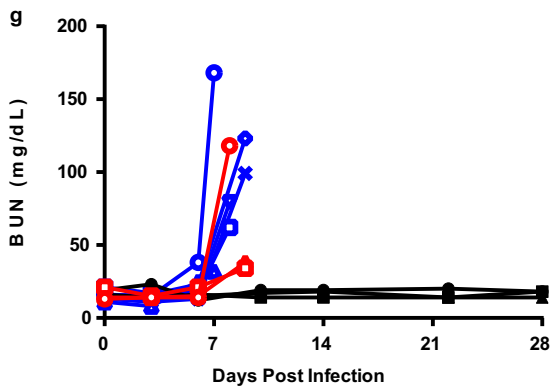
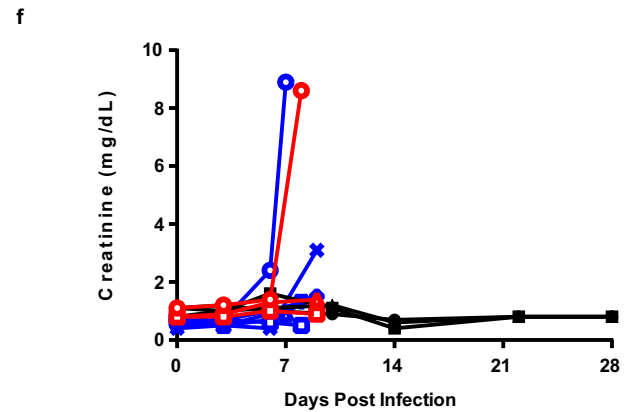
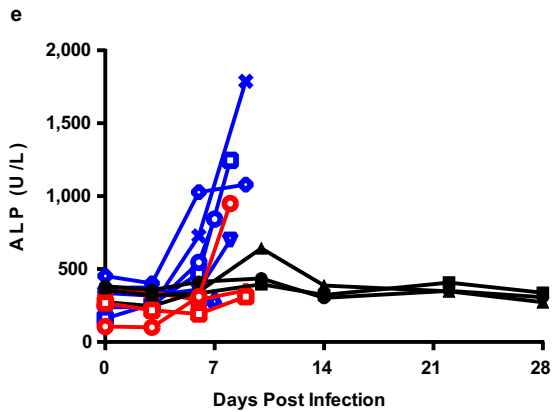
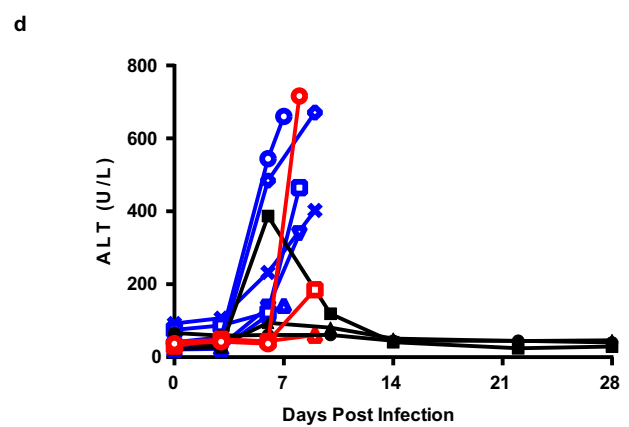
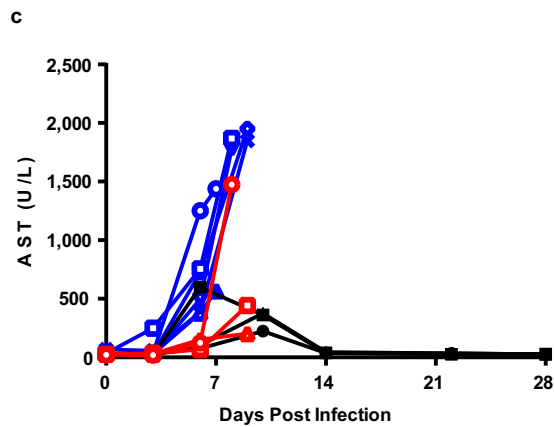
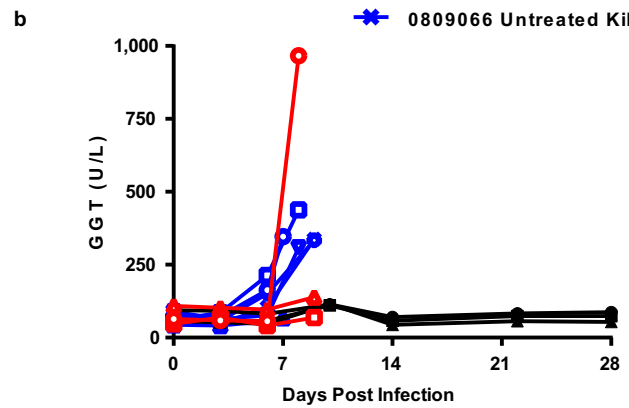
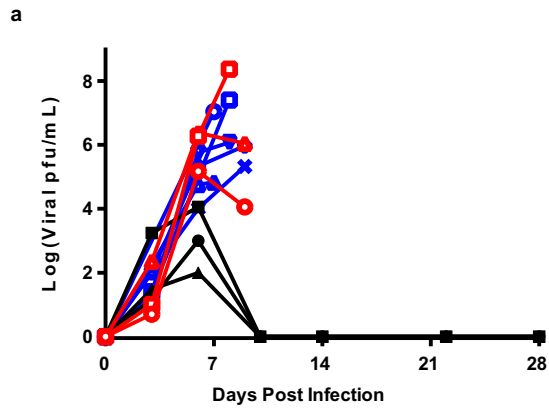


Extended Data Figure 1 | Antiviral activity of siEbola-3 in cells infected with EBOV Makona. For comparison, siEbola-3 activity was also assessed against the central African EBOV Kikwit strain and siEbola-2 activity was evaluated against both EBOV strains. Data are viral RNA copies per millilitre of each sample normalized to untreated infected cells. Results are mean \pm s.e.m. from one biological replicate, conducted in technical triplicate.

0805068 Untreated Makona
1105274 Untreated Makona
JE60 Untreated Makona

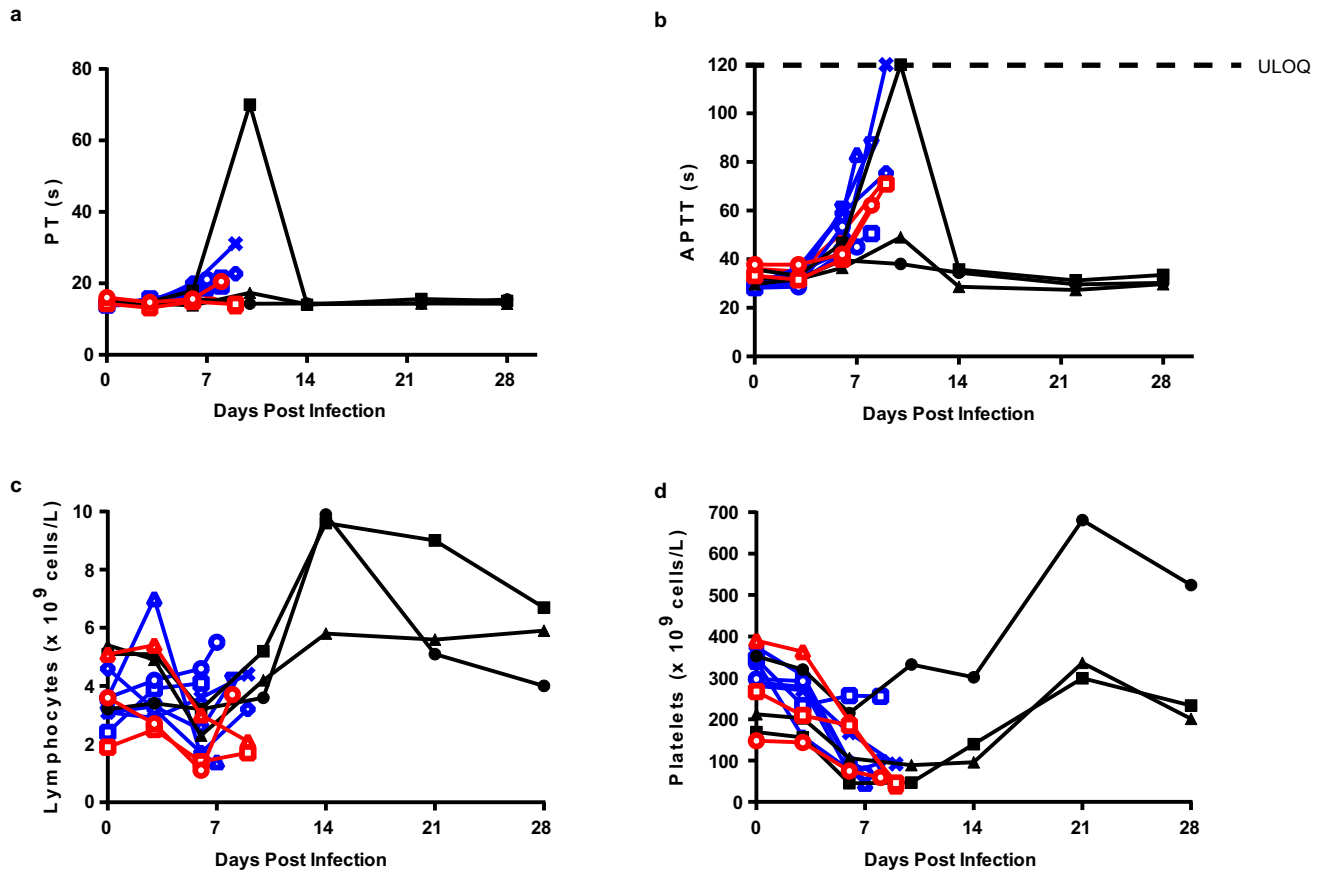
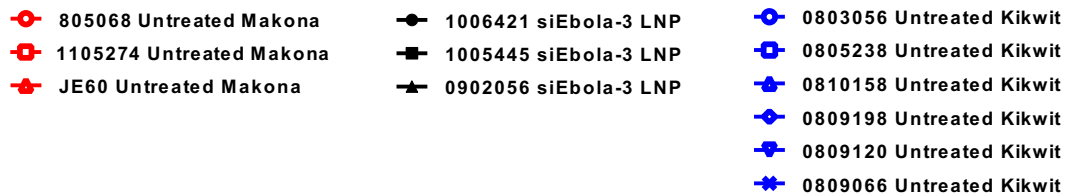
1006421 Makona siEbola-3 LNP Tx
1005445 Makona siEbola-3 LNP Tx
0902056 Makona siEbola-3 LNP Tx

0803056 Untreated Kikwit
0805238 Untreated Kikwit
0810158 Untreated Kikwit
0809198 Untreated Kikwit
0809120 Untreated Kikwit
0809066 Untreated Kikwit



Extended Data Figure 2 | siEbola-3 LNP treatment provides partial protection against EBOV Makona clinical pathologies, and infection with EBOV Makona infection induces a lesser degree of liver dysfunction compared to EBOV Kikwit infection. **a**, No differences in viraemia levels were observed in untreated animals infected with EBOV Makona or Kikwit.

b–e, Liver dysfunction markers. Normal values for uninfected NHPs ranges are GGT ($40\text{--}115\text{ U l}^{-1}$), AST ($20\text{--}45\text{ U l}^{-1}$), ALT ($20\text{--}165\text{ U l}^{-1}$), ALP ($130\text{--}500\text{ U l}^{-1}$). **f, g**, Protection against EBOV-Makona-induced CRE and BUN elevation was observed. Normal values for uninfected NHPs range from BUN ($10\text{--}25\text{ mg dl}^{-1}$) and CRE ($0.8\text{--}1.2\text{ mg dl}^{-1}$).



Extended Data Figure 3 | Comparison of coagulation and haematology characteristics between untreated control animals infected with EBOV Makona or Kikwit. **a, b**, Coagulopathies are not as marked in EBOV Makona

infection when compared to historical EBOV Kikwit data. **c**, Lymphopenia is observed in all infected animals. **d**, Thrombocytopenia levels are similar between EBOV-Makona and EBOV-Kikwit-infected control animals.

Extended Data Table 1 | Comparison of clinical signs progression between untreated rhesus macaques infected with EBOV Makona or EBOV Kikwit

Infection	Animal ID	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9
EBOV Makona	1105274	0	0	0	0	0	0	0	1	16
	0805068	0	0	0	0	0	1	1	15	
	JE60	0	0	0	0	0	1	3	3	14
EBOV Kikwit	809066	0	0	0	0	0	0	0	0	11
	809120	0	0	0	0	0	1	2	17	
	809198	0	0	0	1	1	1	1	3	11
	810158	0	0	0	0	1	1	10		
	805238	0	0	0	0	1	1	14		
	803056	0	0	0	0	0	1	1	10	

Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice

Rene C. Adam¹, Hanseul Yang¹, Shira Rockowitz², Samantha B. Larsen¹, Maria Nikolova¹, Daniel S. Oristian¹, Lisa Polak¹, Meelis Kadaja¹, Amma Asare³, Deyou Zheng^{2,3} & Elaine Fuchs¹

Adult stem cells occur in niches that balance self-renewal with lineage selection and progression during tissue homeostasis. Following injury, culture or transplantation, stem cells outside their niche often display fate flexibility^{1–4}. Here we show that super-enhancers⁵ underlie the identity, lineage commitment and plasticity of adult stem cells *in vivo*. Using hair follicle as a model, we map the global chromatin domains of hair follicle stem cells and their committed progenitors in their native microenvironments. We show that super-enhancers and their dense clusters ('epicentres') of transcription factor binding sites undergo remodelling upon lineage progression. New fate is acquired by decommissioning old and establishing new super-enhancers and/or epicentres, an auto-regulatory process that abates one master regulator subset while enhancing another. We further show that when outside their niche, either *in vitro* or in wound-repair, hair follicle stem cells dynamically remodel super-enhancers in response to changes in their microenvironment. Intriguingly, some key super-enhancers

shift epicentres, enabling their genes to remain active and maintain a transitional state in an ever-changing transcriptional landscape. Finally, we identify SOX9 as a crucial chromatin rheostat of hair follicle stem cell super-enhancers, and provide functional evidence that super-enhancers are dynamic, dense transcription-factor-binding platforms which are acutely sensitive to pioneer master regulators whose levels define not only spatial and temporal features of lineage-status but also stemness, plasticity in transitional states and differentiation.

Hair follicle stem cells fuel cyclical bouts of hair follicle regeneration and hair growth and also repair damaged epidermis⁶. Hair follicle lineage progression is governed in part by dynamic regulation of Polycomb (PcG)-mediated repression/de-repression typified by a trimethylation mark on lysine 27 of histone H3 (H3K27me3)^{7,8}. However, hair follicle stem cell identity and function are mainly independent of PcG-regulated genes, indicating that additional epigenetic mechanisms underlie the governance of critical cell identity genes.

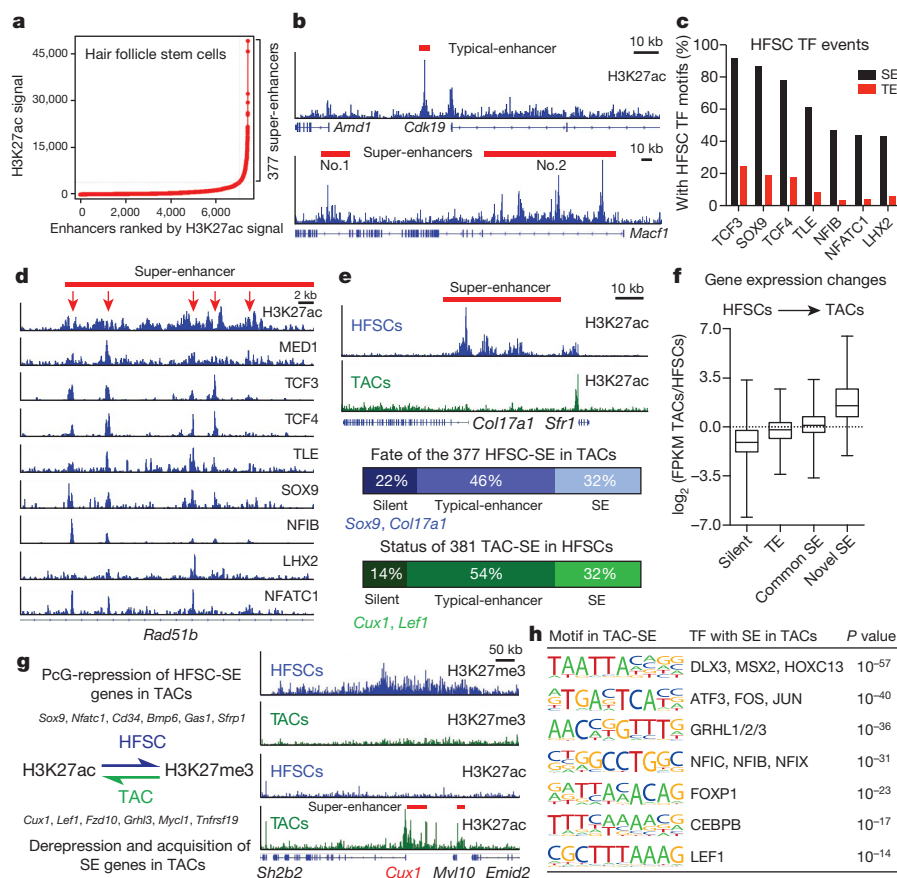


Figure 1 | Dynamic super-enhancer remodelling facilitates lineage progression. **a**, Identification of H3K27ac super-enhancers in hair follicle stem cells. **b**, H3K27ac-marked enhancers at *Cdk19* and *Macf1* loci in hair follicle stem cells. Red bars denote enhancers. **c**, Occupancy of hair follicle stem cell TFs reveals high enrichment particularly within super-enhancers. HFSC, hair follicle stem cell. **d**, Clustering of hair follicle stem cell TFs occurs in epicentres (red arrows) within super-enhancers, as illustrated by the *Rad51b* locus. **e**, Dynamic remodelling of super-enhancers during lineage progression. Example shows that the super-enhancer of the *Col17a1* locus, active in hair follicle stem cells, undergoes complete decommissioning in TACs. **f**, Enhancer remodelling correlates with gene expression changes. Boxplot displaying the full range of variation (min. to max.) shows that the changes that take place in super-enhancers have a profound effect on gene expression during lineage progression ($n = 2$). **g**, Cell fate determinants switch between super-enhancer activation and PcG repression, typified by a swap in H3K27 modifications. Representative example (*Cux1*) shows this switch upon hair follicle stem cell to TAC fate commitment. **h**, Motif analysis of TAC super-enhancers for putative TF binding sites reveals that they are regulated by a new cohort of transcriptional regulators. SE, super-enhancer; TE, typical-enhancer.

¹Howard Hughes Medical Institute, Laboratory of Mammalian Cell Biology & Development, The Rockefeller University, New York, New York 10065, USA. ²Department of Genetics, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ³Departments of Neurology and Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461, USA.

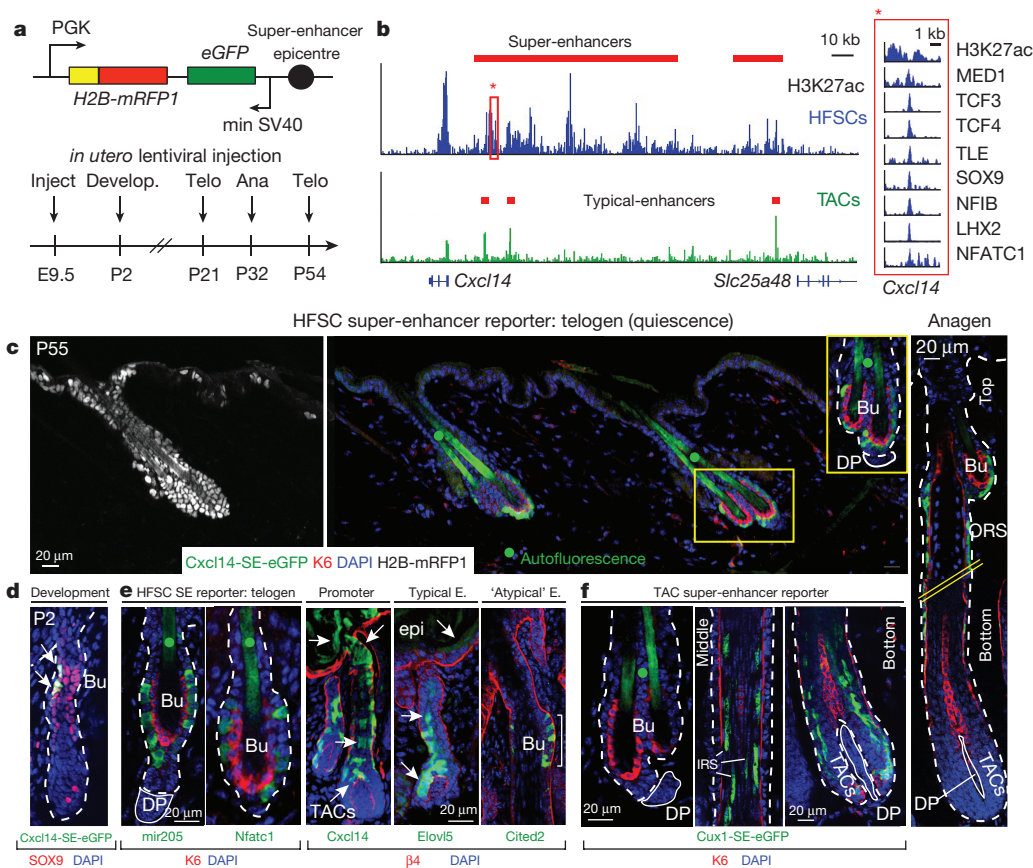


Figure 2 | Super-enhancer epicentres confer tissue, lineage and temporal specificity. **a**, Lentiviral super-enhancer reporter and analysis scheme. Telo: telogen (quiescent hair follicle stem cells; no TACs, no hair growth); Ana: anagen (active TACs, hair growth). **b**, H3K27ac occupancy at the *Cxcl14* locus. Red box highlights the *Cxcl14* super-enhancer epicentre (bound by MED1 and seven hair follicle stem cell TFs; absent in TACs) cloned for reporter assays. **c**, *Cxcl14-SE-eGFP* expression in H2B-mRFP⁺ epidermis is limited to hair follicle stem cells and early hair follicle stem cell progeny along upper ORS in anagen (right). Hatched lines denote the spliced out middle-region of the hair follicle. **d**, Temporal activation of *Cxcl14-SE-eGFP* in SOX9⁺ cells,

Recent *in vitro* studies suggest that genes controlling unique cellular identities are driven by so-called ‘super-enhancers’^{25,9,10}. Representing a small fraction of total enhancers, super-enhancers encompass large chromatin domains bountiful in cell-type specific transcription factor (TF) binding motifs that enable TFs to bind cooperatively. Their richness in H3K27 acetylation renders super-enhancers mutually exclusive for H3K27me3 repression^{5,11–13}, while their H3K4me1 and Mediator complex alliances facilitate interactions with promoters to initiate transcription¹⁴.

To explore the *in vivo* importance of super-enhancers in stem cells, we first conducted chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq) on hair follicle stem cells purified directly from skin (Extended Data Fig. 1). H3K27ac, Mediator subunit MED1 and H3K4me1 peaks resided within promoters (\pm 2 kb of annotated genes) (40%) and distal elements, considered enhancers (60%) of hair follicle stem cell chromatin. A total of 377 super-enhancers were identified by size (>28 kb) and elevated H3K27ac occupancy⁷ with ≥ 5 H3K27ac-enriched clusters (Fig. 1a, b and Extended Data Fig. 2a-f).

A total of >80% accuracy in super-enhancer gene assignments can be achieved by applying optimized RNA-seq and proximity algorithms¹⁴. Most remaining ambiguities arise from multiple expressed genes in close proximity of a super-enhancer¹⁴. We resolved these by requiring that hair follicle stem cell super-enhancer genes must (1) exhibit H3K4me3/H3K79me2-activating and lack H3K27me3-repressive modifications⁸; and (2) maintain strict correlation between super-enhancer and

concomitant with hair follicle stem cell niche establishment at P2. **e**, Hair-follicle-stem-cell-specific targeting by *mir205*- and *Nfatc1*-super-enhancer epicentres, whereas the *Cxcl14* promoter and *Elovl5* typical-enhancer display broader activity. Atypical *Cited2*-TE binds all seven hair follicle stem cell TFs and drives hair-follicle-stem-cell-specific targeting. **f**, *Cux1*-SE-eGFP is silent in hair follicle stem cells, but activated during the hair cycle in TACs and differentiating IRS progeny. Dotted lines denote epidermal–dermal border; solid lines delineate DP (dermal papilla). Bu, bulge (hair follicle stem cell niche). Green dot denotes hair shaft autofluorescence.

candidate expression in three different states: hair follicle stem cells, their committed progenitors *in vivo*, and hair follicle stem cells *in vitro* (Supplementary Table 1; see below).

Whereas typical-enhancers (1–2 kb) governed >90% of hair follicle stem cell genes, super-enhancers marked genes transcribed selectively in hair follicle stem cells (Extended Data Fig. 2g, h). Unbiased gene ontology (GO) analysis further distinguished super-enhancer regulated genes by a preponderance of transcriptional regulators, including *Sox9*, *Lhx2*, *Nfatc1* and *Nfib*, important for stemness, quiescence and/or crosstalk within the hair follicle stem cell niche^{15–18} (Extended Data Fig. 2i, j). Their encoded TFs, in particular SOX9, bound at high frequency (87%) to super-enhancers, including their own, indicative of auto-regulation (Fig. 1c). Essential hair follicle stem cell WNT-effector TCF3 (ref. 19) also bound within these super-enhancers, although the *Tcf7l1* enhancer fell just below our assignment cut-off.

Notably, >60% of super-enhancers were occupied by ≥ 5 different hair follicle stem cell TFs. The hair follicle stem cell TF binding was not similarly distributed within open chromatin of comparable cohorts of typical-enhancers, even when flanking sequences were included to normalize for their smaller size (Extended Data Fig. 3a, b). Thus, binding of hair-follicle-stem-cell-specific TFs was not dictated by open chromatin *per se*, but rather by super-enhancers, which controlled critical cell identity genes, including themselves, in this adult stem cell niche.

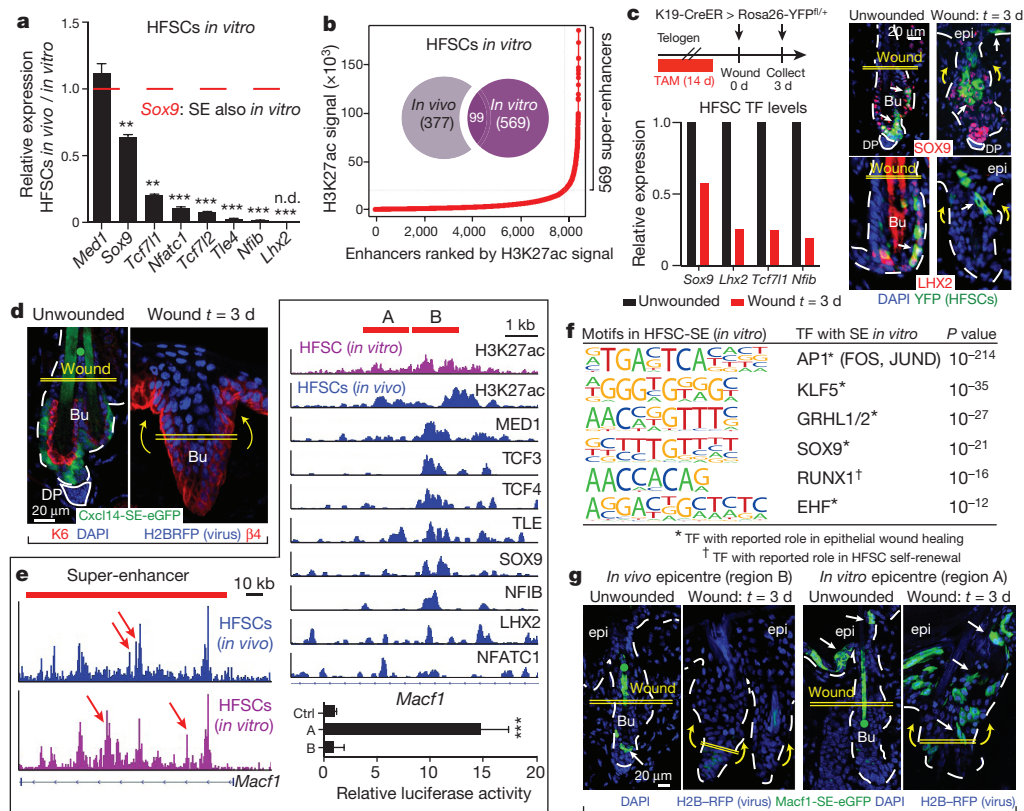


Figure 3 | The sensitivity of super-enhancers to environmental changes allows hair follicle stem cell adaptation and plasticity. **a**, Repression of hair follicle stem cell TF genes *in vitro*. Mean and standard deviation are shown ($n = 3$). *P* values from *t*-test: ** $P < 0.01$; *** $P < 0.001$; n.d., not detected. **b**, Super-enhancers in hair follicle stem cells show little overlap *in vivo* and *in vitro*. **c**, Downregulation of hair follicle stem cell TFs during wound-repair in K19-CreER(hair-follicle-stem-cell-specific)/R26YFP mice. SOX9 is present in migrating hair follicle stem cells at reduced levels. **d**, *Cxcl14-SE-eGFP* reporter is repressed in wound-induced hair follicle stem cells. **e**, *In vivo* versus

Scattered across each super-enhancer were smaller (1–2 kb) regions densely packed with hair follicle stem cell TF consensus binding motifs and which bound the cohort of hair follicle stem cell TFs (Fig. 1d). These epicentres resembled recently described ‘hotspots’ within super-enhancers of cultured adipocytes²⁰. Notably, <1% of typical-enhancers had even one such cluster of hair follicle stem cell TF motifs, whereas most hair follicle stem cell super-enhancers had ten (Extended Data Fig. 3).

An auto-regulatory and cooperative mechanism⁵ predicts that super-enhancer remodelling must occur to progress along a lineage typified by environmentally induced changes in TF landscape. We tested this hypothesis by characterizing the super-enhancers of short-lived hair follicle stem cell progeny (transit-amplifying cells, TACs) that progress to make hair (Extended Data Fig. 1). The 381 super-enhancer-marked TAC genes diverged considerably from those of hair follicle stem cells (Fig. 1e). Notably, hair follicle stem cell TF genes lost their super-enhancers in TACs, while TAC TF genes gained super-enhancers. Thus, our findings broadened the concept of super-enhancer dynamics observed in macrophages isolated from different tissues^{11,12}, and supported the notion that enhancers are activated or silenced in lineage-specific fashion^{8,21}. However, they contrasted with prior *in vivo* studies suggesting that chromatin remains broadly permissive as intestinal stem cells progress through a lineage²².

Like hair follicle stem cells, TAC super-enhancers controlled TF, BMP and WNT signalling genes, but the presence of cell-cycle related and NOTCH pathway super-enhancer-marked genes appeared unique to features of TACs (Extended Data Fig. 4). Interestingly, only 32% of hair follicle stem cell super-enhancers persisted in TACs. Half were reduced

in vitro hair follicle stem cell differences in H3K27ac peak (epicentre) distributions (arrows) of the super-enhancer of *Macf1*. One epicentre shift is magnified at right. Region B represents an epicentre active *in vivo*, richly bound by hair follicle stem cell TFs; adjacent region A is an epicentre active *in vitro*. Mean and standard deviation of relative luciferase activities are shown below ($n = 3$). *P* value from *t*-test: *** $P < 0.001$. **f**, Motif analysis for TF binding sites of *in vitro* hair follicle stem cell super-enhancers. **g**, Functional validation of epicentre shifts in mice transduced with *Macf1-SE-eGFP* reporters. Note dynamic changes in reporter activity upon wounding.

to typical-enhancers, suggestive of more subordinate roles. Analogously, 54% of genes that gained a super-enhancer in TACs were driven by typical-enhancers in hair follicle stem cells (Fig. 1e). Typical-enhancer to super-enhancer shifts correlated with increased transcription and appeared to provide an epigenetic readout to gauge transcriptional levels during lineage progression (Fig. 1f).

Most super-enhancer genes involved in dictating hair follicle stem cell fate were decommissioned in TACs. For this cohort, H3K27ac loss was accompanied by H3K27me3 gain⁸, suggestive of ‘super-silencing’ (Fig. 1g). Conversely, specific TAC fate determinants became de-repressed by losing PcG-catalysed H3K27me3 marks, while simultaneously gaining H3K27ac to expose a new super-enhancer. An unbiased analysis revealed that TAC super-enhancers were enriched for the binding motifs of many TAC TFs (Fig. 1h).

To address functionality, we tested the ability of super-enhancer epicentres to drive reporter gene expression *in vivo*. A 1.2 kb epicentre within the *Cxcl14* super-enhancer was used to generate a high-titre lentivirus harbouring *Cxcl14-SE-eGFP* (SE, super-enhancer) and *Pgk-H2B-mRFP1* and injected into the amniotic sacs of living E9.5 embryos (Fig. 2a, b). This results in random transgene integration into skin progenitor chromatin²³.

By the adult stage, H2B-mRFP1 was expressed throughout skin epithelium. Notably, however, eGFP was confined to hair follicle stem cells that reside in the outer layer of the resting (telogen) phase ‘bulge’ niche (Fig. 2c and Extended Data Fig. 5). As a new hair cycle began, *Cxcl14-SE-eGFP* activity persisted in hair follicle stem cells and early transitory progeny within the upper outer root sheath (ORS) of regenerating hair

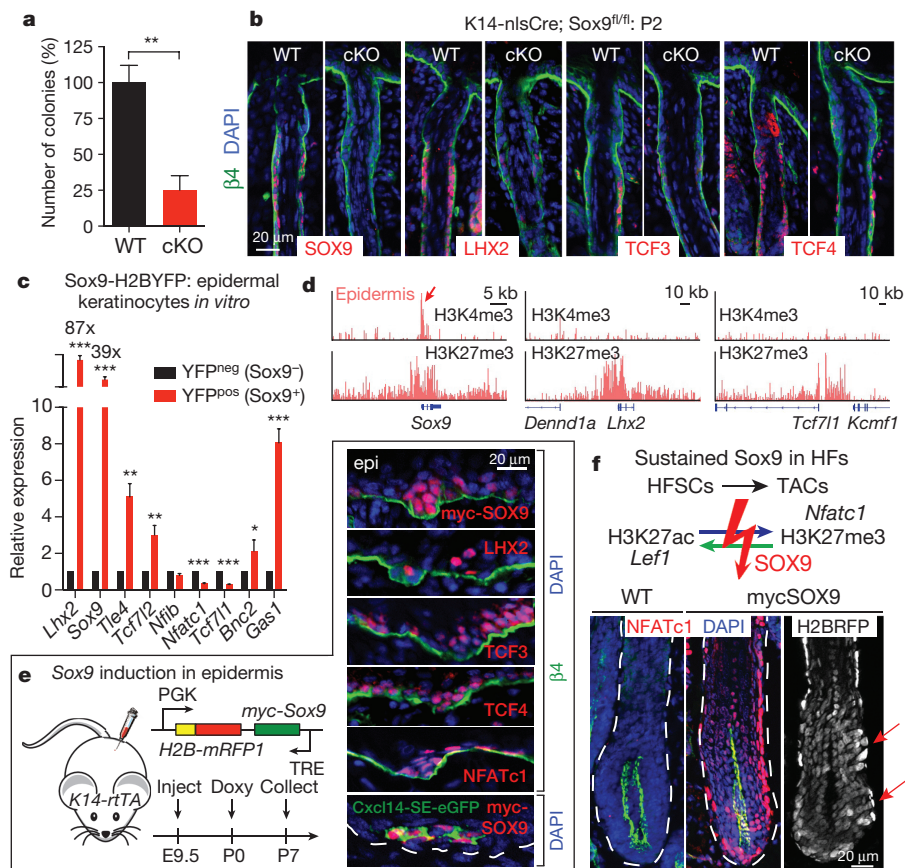


Figure 4 | SOX9 is a pioneer factor governing hair follicle stem cell fate and plasticity. **a**, Colony formation assays on wild-type (WT) and *Sox9*-cKO hair follicle stem cells. Mean and standard deviation are shown ($n = 3$). P value from t -test: $**P < 0.01$. **b**, Hair follicle stem cell specification fails in *Sox9*-cKO mice. **c**, Ectopic *Sox9* in epidermal keratinocytes induces hair follicle stem cell super-enhancer genes. Mean and standard deviation are shown ($n = 3$). P values from t -test: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. **d**, Hair follicle stem cell TF genes *Lhx2* and *Tcf7l1* are PcG-repressed in epidermis *in vivo*, while *Sox9* is poised. **e**, Forced *in vivo* expression of *Sox9* in epidermal progenitors activates hair follicle stem cell regulated *Cxcl14*-SE-eGFP. **f**, Sustained SOX9 during hair follicle regeneration. Note prevention of swap in H3K27 modifications of key super-enhancers upon TAC fate commitment. NFATc1 atypically persists in lower ORS and TACs, and minimilge-like structures occur along ORS (arrows).

follicles. By contrast, the reporter was silenced in committed TACs, which lack hair follicle stem cell TFs altogether. The specificity of *Cxcl14*-SE-eGFP extended to development, where it became faithfully activated coincident with hair follicle stem cell TF and niche establishment²⁴ (Fig. 2d).

Similar eGFP patterns were observed when hair follicle stem cell super-enhancer epicentres from *Nfatc1* and *mir205* were used as drivers. However, in contrast to its super-enhancer, the promoter of *Cxcl14* drove broad skin epithelial expression, as did the typical-enhancer of *Elovl5*. Despite some tissue-specific elements, these regions lacked clustered hair follicle stem cell TF binding sites. By contrast, the typical-enhancer of *Cited2* uncharacteristically contained all hair follicle stem cell TF motifs and correspondingly exhibited hair follicle stem cell specificity. Summarized in Fig. 2e, these results provide compelling evidence that concentration of binding sites for a diverse array of hair follicle stem cell TFs is what confers lineage and stage-restricted specificity, whose activity is largely refractory to integration site.

Prior knowledge of master regulators was not necessary to tease apart these specialized regulatory elements from those driving broader expression. Thus, by identifying a MED1-bound, H3K27ac-intense epicentre of the TAC super-enhancer-controlled *Cux1* gene²⁵, we could generate a reporter with activity restricted to the hair follicle channel (inner root sheath, IRS) TACs (Fig. 2f and Extended Data Fig. 5). These findings illustrate the power of super-enhancers and their epicentres for developing genetic tools with unprecedented cell-type, temporal, lineage and stage specificity.

Intriguingly, *Cxcl14*-SE-eGFP activity was silent in cultured hair follicle stem cells, consistent with the complete decommissioning of hair follicle stem cell niche super-enhancers *in vitro* (Extended Data Fig. 6). Super-enhancer-associated hair follicle stem cell TF genes were also repressed *in vitro*, but upon engraftment²⁶ were faithfully restored. This behaviour suggested that super-enhancer epicentres are reversibly sensitive to their microenvironment. Additionally, although new super-enhancers were acquired *in vitro*, few corresponded to 'signature genes'

of hair follicle stem cells proliferating *in vivo*⁶, suggesting environmental adaptation is more critical than proliferative status (Fig. 3a, b and Extended Data Fig. 6).

Curiously, some genes acquiring super-enhancers *in vitro* were epidermal genes, while others have been implicated in wound-repair. If these dynamics reflect a transitional state analogous to early stages of wound-repair, super-enhancer regulated genes should change quickly once hair follicle stem cells exit their niche and migrate into a wound bed. To test this possibility, we fluorescently tagged niche hair follicle stem cells, introduced a shallow wound, and monitored for proteins whose genes had changed super-enhancer status *in vitro*. Shortly after epidermal injury, YFP-marked hair follicle stem cells downregulated both super-enhancer-regulated TF genes and LV-transduced *Cxcl14*-SE-eGFP (Fig. 3c, d). Conversely, wound-activated hair follicle stem cell progeny induced *Fhl2* and *Prrg4*, which in culture, displayed super-enhancer-mediated activity. Moreover, upon transplantation these *in vitro* induced genes were silenced concomitant with hair follicle regeneration (Extended Data Fig. 6j). Together, these findings underscore the sensitivity of super-enhancers to their microenvironment^{11,12}, and directly link the relevance of culture-induced super-enhancer dynamics to wound-repair and fate plasticity.

A small cohort of genes maintained super-enhancers *in vitro*, including those from a recent hair follicle stem cell self-renewal screen²⁷ and genes like *Macf1* that function in wound-repair²⁸. Seeking to determine how these super-enhancers remain active in the face of downregulated hair follicle stem cell TFs, we noticed that their epicentres had shifted in culture (Fig. 3e). Upon analysing several of these shifts, we discovered that instead of hair follicle stem cell TF motifs, *in vitro* epicentres were enriched for epidermal/wound-related motifs. These included AP1, KLF, grainyhead-like and FOX families, many of whose genes displayed *in vitro* specific super-enhancers (Fig. 3f and Extended Data Fig. 7a–c).

Using reporter assays, we tested functionality of seven different epicentre shifts. *In vitro*, physiological hair follicle stem cell epicentres

exhibited no luciferase activity, while culture-based epicentres were robustly active. Conversely, when tested *in vivo*, *in vitro* epicentres were only active in epidermis, while physiological hair follicle stem cell epicentres restricted expression to the hair follicle niche. However following injury, *in vitro* epicentres were induced in activated hair follicle cells undergoing epidermal repair, while epicentres of quiescent hair follicle stem cells became repressed (Fig. 3g).

While these results are intriguing, it remains unclear how hair follicle stem cells exploit super-enhancer dynamics to elicit the plasticity that allows them to regenerate hair follicles during homeostasis, repair damaged epidermis following injury, and adapt to culture. We were drawn to SOX9, as *Sox9* was the only hair follicle stem cell TF gene that maintained a super-enhancer *in vitro*, where it was expressed at lower levels (Fig. 3a and Extended Data Fig. 8). It was also maintained at reduced levels in wound-activated hair follicle stem cells, suggesting a role for SOX9 in transitional states (Fig. 3c). Curiously, quantitative *Sox9* ablation in adult hair follicle stem cells¹⁶ strikingly reduced colony-forming efficiency *in vitro*, and *Sox9* ablation during skin embryogenesis blocked LHX2, TCF3 and TCF4 expression and formation of functional hair follicle stem cells (Fig. 4a, b).

Conversely, ectopic SOX9 in cultured hair follicle stem cells induced *Lhx2*, *Tcf7l1* and *Tcf7l2* transcription. Even more impressive were the effects of SOX9 on epidermal keratinocytes, which *in vitro* as *in vivo*, did not express hair follicle stem cell TFs. *Lhx2* showed >80× elevation upon SOX9 induction in epidermal cultures and repression shortly after *Sox9* ablation in hair follicle stem cells. Neither *Sox9*, *Tcf7l1* nor *Tcf7l2* showed such sensitivity to LHX2, indicating a special importance for SOX9 in regulating hair follicle stem cell super-enhancer activity (Fig. 4c and Extended Data Fig. 8g).

If SOX9 is a true pioneer factor whose levels dictate whether super-enhancers will be epigenetically active or silenced, then inducing SOX9 in skin epidermis should activate genes such as *Tcf7l1* and *Lhx2* whose super-enhancers are PcG-silenced (Fig. 4d). We tested this possibility with a doxycycline-inducible SOX9 lentivirus transduced *in utero* into K14-rtTA animals and activated at P0 or P21. Notably, SOX9 expression in epidermis activated other super-enhancer controlled hair follicle stem cell TF genes. The ability of SOX9 to initiate H3K27 acetylation was exemplified by its activation of normally PcG-silenced *Lhx2* and *Tcf7l1* in epidermis. The activation of *Cxcl14-SE-eGFP* in SOX9-expressing epidermis explicitly traced the phenomenon to super-enhancers (Fig. 4e).

Finally, prolonging SOX9 in the hair follicle lineage generated equally striking perturbations. The lower ORS was riddled with mini-bulge-like structures concomitantly with persistent LHX2, TCF3/4 and SOX9 in this transitory zone. NFATc1 was atypically sustained in lower ORS and TACs, while the switch to TAC super-enhancers was impaired (Fig. 4f and Extended Data Fig. 9). The failure of *Nfatc1* to become PcG-silenced in SOX9⁺ TACs shows that SOX9 protects against H3K27me3 silencing at super-enhancer regulated genes. In summary, by coupling a pioneer factor, SOX9, which senses local changes in microenvironment, to chromatin platforms optimized for sensing TF concentration, super-enhancers elicit the chromatin dynamics required for skin stem cells to pursue distinct lineages, repair wounds and exhibit plasticity in transitional states.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 September 2014; accepted 6 February 2015.

Published online 18 March; corrected online 20 May 2015 (see full-text

HTML version for details).

1. Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414**, 98–104 (2001).
2. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).

3. van Es, J. H. *et al.* Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biol.* **14**, 1099–1104 (2012).
4. Blanpain, C. & Fuchs, E. Stem cell plasticity. Plasticity of epithelial stem cells in tissue regeneration. *Science* **344**, 1242281 (2014).
5. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
6. Hsu, Y. C., Li, L. & Fuchs, E. Transit-amplifying cells orchestrate stem cell activity and tissue regeneration. *Cell* **157**, 935–949 (2014).
7. Ezhkova, E. *et al.* EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev.* **25**, 485–498 (2011).
8. Lien, W. H. *et al.* Genome-wide maps of histone modifications unwind *in vivo* chromatin states of the hair follicle lineage. *Cell Stem Cell* **9**, 219–232 (2011).
9. Loven, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
10. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
11. Gosselin, D. *et al.* Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell* **159**, 1327–1340 (2014).
12. Lavin, Y. *et al.* Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* **159**, 1312–1326 (2014).
13. Liu, Z. *et al.* Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell* **159**, 358–373 (2014).
14. Downen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
15. Folgueras, A. R. *et al.* Architectural niche organization by LHX2 is linked to hair follicle stem cell function. *Cell Stem Cell* **13**, 314–327 (2013).
16. Kadaja, M. *et al.* SOX9: a stem cell transcriptional regulator of secreted niche signaling factors. *Genes Dev.* **28**, 328–341 (2014).
17. Keyes, B. E. *et al.* Nfatc1 orchestrates aging in hair follicle stem cells. *Proc Natl Acad Sci USA* **110**, E4950–E4959 (2013).
18. Chang, C. Y. *et al.* NFIB is a governor of epithelial-melanocyte stem cell behaviour in a shared niche. *Nature* **495**, 98–102 (2013).
19. Lien, W. H. *et al.* *In vivo* transcriptional governance of hair follicle stem cells by canonical Wnt regulators. *Nature Cell Biol.* **16**, 179–190 (2014).
20. Siersbaek, R. *et al.* Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Rep.* **7**, 1443–1455 (2014).
21. Luyten, A., Zang, C., Liu, X. S. & Shivdasani, R. A. Active enhancers are delineated *de novo* during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev.* **28**, 1827–1839 (2014).
22. Kim, T. H. *et al.* Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity. *Nature* **506**, 511–515 (2014).
23. Beronja, S., Livshits, G., Williams, S. & Fuchs, E. Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nature Med.* **16**, 821–827 (2010).
24. Nowak, J. A., Polak, L., Pasolli, H. A. & Fuchs, E. Hair follicle stem cells are specified and function in early skin morphogenesis. *Cell Stem Cell* **3**, 33–43 (2008).
25. Ellis, T. *et al.* The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. *Genes Dev.* **15**, 2307–2319 (2001).
26. Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. & Fuchs, E. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635–648 (2004).
27. Chen, T. *et al.* An RNA interference screen uncovers a new molecule in stem cell self-renewal and long-term regeneration. *Nature* **485**, 104–108 (2012).
28. Wu, X. *et al.* Skin stem cells orchestrate directional migration by regulating microtubule-ACF7 connections through GSK3β. *Cell* **144**, 341–352 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Mazel, L. Li, S. Semova, and S. Tadesse for FACS sorting (Rockefeller University FACS facility); and C. Lai for assistance in high-throughput sequencing (Rockefeller University Genomics Resource Center). We thank E.F. laboratory members A. Aldegue, S. Hacker, M. Sribour and J. Levorse for assistance in mouse research; I. Matos for advice on image acquisition; J. Racelis, S. Chai, and E. Wong for technical assistance; and Y. Ge, S. Naik, A. Kulukian, and N. Oshimori for discussions. R.C.A. was supported by the Anderson Cancer Center Graduate Student Fellowship. E.F. is an HHMI Investigator. This work was supported by grants from the National Institutes of Health (R01-AR31737 to E.F. and R21MH099452 to D.Z.).

Author Contributions R.C.A. and E.F. conceived the project and designed the experiments. R.C.A. and H.Y. performed the experiments, including FACS purification, ChIP-seq assays, data analysis and *in vivo* reporter assays. D.S.O. and L.P. carried out *in utero* lentiviral injections and wounding experiments with mice. S.B.L., M.N. and M.K. contributed to *in vitro* experiments. A.A., S.R. and D.Z. performed bioinformatics analyses. E.F. supervised the project. R.C.A. and E.F. wrote the manuscript.

Author Information ChIP-seq data are deposited in GEO under accession number GSE61316. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.F. (fuchslb@rockefeller.edu).

METHODS

Mouse lines. Female CD1 mice (8 weeks old, Charles River) were used for the purification of hair follicle stem cells. Female CD-1 mice transgenic for Krt14-H2B-GFP²⁹ (30–32 days old) were used for the purification of TACs. Krt15-CrePGR; Sox9^{fl/n}; R26YFP^{fl/+} mice have been described¹⁶. Krt19-CreER mice have been described³⁰. CreER was activated by intraperitoneal injection of mice with 20 mg ml⁻¹ tamoxifen (Sigma) in corn oil (Sigma) to specifically label hair follicle stem cells. For the generation of K14-H2B-iRFP mice, iRFP was first amplified from pShuttle-CMV-iRFP (Addgene plasmid 31856) and fused with H2B, before the H2B-iRFP construct was assembled with the Krt14 promoter, β -globin intron and poly(A) sequences³¹. Transgenic mice were generated with standard pronuclear injections. For lentiviral injections, transduced mice were confirmed by genotyping with RFP primers: forward 5'-ATCCTGTCCCTCAGTTCAGTAC-3', reverse 5'-TCCACGATGGTGTAGTCTCTCGTTG-3'. For TRE-mycSox9 transduced mice, positive mice were fed with doxycycline-containing chow, starting either at P0 (newborn) or at P21 (adult). No formal randomization was performed, and studies were not blinded. Mice were maintained in the Association for Assessment and Accreditation of Laboratory Animal Care-accredited animal facility of The Rockefeller University (RU), and procedures were performed with Institutional Animal Care and Use Committee (IACUC)-approved protocols.

Flow cytometry. Preparation of adult mice back skins for isolation of hair follicle stem cells and TACs were done as previously described^{8,32}. Briefly, for telogen skin, subcutaneous fat was removed with a scalpel, and skins were placed dermis side down on trypsin (Gibco) at 37 °C for 35 min. Single-cell suspensions were obtained by scraping the skin gently. Anagen skin was treated with collagenase at 37 °C for 30 min to dissociate dermal cells and then incubated with trypsin at 37 °C for 15 min to detach and generate single-cell suspensions of the epidermal and HF cells. Cells were then washed with PBS containing 5% of fetal bovine serum (FBS), then filtered through 70 μ m and 40 μ m cell strainers. Cell suspensions were incubated with the appropriate antibodies for 30 min on ice. The following antibodies were used for FACS: α 6-PE (1:100, eBiosciences), CD34-eFluor660 (1:100, eBiosciences) and Sca-1-PerCP-Cy5.5 (1:1,000, eBiosciences). DAPI was used to exclude dead cells. Cell isolations were performed on FACSria sorters running FACSDiva software (BD Biosciences).

ChIP-seq. Immunoprecipitations were performed on FACS-sorted populations from female mice or on cultured hair follicle stem cells⁸. For each ChIP-seq run, 7×10^6 to 2×10^7 cells were used. Antibodies used for ChIP-seq were anti-H3K27ac (abcam, ab4729), anti-H3K4me1 (abcam, ab8895), anti-Crspl/Trap220 (Med1, Bethyl Laboratories, A300-793A) and anti-H3K27me3 (Millipore, 07-449). Briefly, cells were cross-linked in 1% (wt/vol) formaldehyde solution, resuspended, and lysed. To solubilize and shear cross-linked DNAs, lysates were subjected to a Bio-ruptor Sonicator (Diagenode, UCD-200) according to a 30 \times regimen of 30 s sonication followed by 60 s rest. The resulting whole-cell extract was incubated overnight at 4 °C with 10 μ l of Dynabeads Protein G magnetic beads (Life Technologies) which had been pre-incubated with 5 μ g of the appropriate Ab. After ChIP, samples were washed, and bound complexes were eluted and reverse cross-linked. ChIP DNA was prepared for sequencing by repairing sheared DNA and adding Adaptor Oligo Mix (Illumina) in the ligation step. A subsequent PCR step with 25 amplification cycles added the additional Solexa linker sequence to the fragments to prepare them for annealing to the Genome Analyzer flow cell. After amplification, a range of fragment sizes between 150–300 bp was selected and the DNA was gel-purified and diluted to 10 nM for loading on the flow cell. Sequencing was performed on the Illumina HiSeq 2500 Sequencer following manufacturer protocols. ChIP-seq reads were aligned to the mouse genome (mm9, build 37) using Bowtie aligner³³. ChIP-seq signal tracks were presented by Integrative Genomics Viewer (IGV) software.

Bioinformatics analysis. H3K27ac peaks were called by the program MACS³⁴ (v1.4.2, default parameters) from the ChIP-seq data with the input as controls. The peaks were associated to genes using the mouse RefSeq annotations; those located within 2 kb of transcription start sites were called as 'promoter' peaks and the rest were 'enhancer' peaks. The H3K27ac enhancer peaks were used for the identification of super-enhancers, using the algorithm described previously, wherein enhancer peaks were stitched together if they are located within 12.5 kb of each other and if they don't have multiple active promoters in between. Enhancers were then ranked according to increasing H3K27ac signal intensity⁷. Enhancer-gene assignments were performed using the following criteria to make gene assignments: (1) proximity of genes to the super-enhancer of stem cells; (2) high transcriptional activity in stem cells (by RNA-seq and by ChIP-seq for presence of H3K4me3/H3K79me2 marks and no H3K27me3 marks in the promoter/typical-enhancer and/or gene body); (3) correlation between loss of the super-enhancer (or shift in its epicentre peaks), loss of gene transcription and loss of H3K79me2 mark \pm H3K27me3 mark in proliferative short-lived progenitors; (4) correlation between loss of the super-enhancer (or shift in its epicentre), loss of gene transcription and loss of H3K79me2 mark \pm H3K27me3 mark in proliferative cultured stem cells. The overlap of super-enhancers

with ChIP-seq peaks for MED1 and other TFs was defined by ≥ 1 base overlap. For TF enrichment analysis at super-enhancers, H3K27ac peaks not located at super-enhancers (that is, typical-enhancers) were randomly picked, extended to match the sizes of super-enhancers and used as background controls. GO function enrichment analyses were carried out by the software GREAT³⁵ using the list of super-enhancer coordinates and the default setting. For motif analysis of enhancers located in super-enhancers, 1-kb sequences under the H3K27ac peaks were searched for enriched motifs using the software HOMER (v4.6) with the default setting (PMID 20513432). Epicentres were defined as 1-kb-regions flanking either side of the H3K27ac peaks. The 1-kb size was chosen based on our analysis of the distances of H3K27ac peaks to their nearest transcription factor ChIP-seq peaks in hair follicle stem cells *in vivo* (distance of the two peak centres, Extended Data Fig. 3e), which showed an enrichment of TF binding within 1-kb regions of H3K27ac peaks. Overlapping epicentres were merged during this analysis. To analyse epicentre shifting, for each of the overlapping super-enhancers between hair follicle stem cells *in vivo* and *in vitro*, we determined the number of epicentres that were not overlapping in the two samples and considered them as shifting epicentres. To generate the heatmap (Extended Data Fig. 3c), the program seqMiner³⁶ was used to calculate the ChIP-seq read densities, which were the maximal numbers of overlapping ChIP-seq reads in 50-bp bins from -5 kb to +5 kb of the H3K27ac peak summits. The density matrix was clustered based on the H3K27ac ChIP-signal and then used to generate a heatmap.

Antibodies. The following antibodies and dilutions were used: SOX9 (rabbit, 1:1,000, Millipore), NFIB (rabbit, 1:1,000, Active Motif), LHX2 (rabbit, 1:2,000, Fuchs lab), K6 (guinea pig, 1:5,000, Fuchs laboratory), K24 (rabbit, 1:5,000, Fuchs lab), CD34 (rat, 1:100, BD-Pharmingen), LEF1 (rabbit, 1:100, Fuchs lab), NFATc1 (mouse, 1:100, Santa Cruz), TCF3 (guinea pig, 1:200; Fuchs laboratory), TCF4 (rabbit, 1:300; Cell Signaling Technology), FHL2 (rabbit, 1:100, Abcam), PRRG4 (rabbit, 1:100, Abcam), CUX1 (rabbit, 1:200, Santa Cruz), β 4-integrin (rat, 1:100, BD-Pharmingen), GFP (chicken, 1:2,000, Abcam), RFP (rabbit, 1:5,000, MBL; or guinea-pig, 1:3000, Fuchs laboratory). Secondary antibodies coupled to Alexa488, RRX, or Alexa647 were from Life Technologies. Nuclei were stained using 4',6'-diamidino-2-phenylindole (DAPI).

Histology, immunofluorescence and imaging. Back skins from mice were embedded in OCT (Tissue Tek), frozen, cryosectioned (10–20 μ m) and fixed for 10 min in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS). For lentivirally transduced mice, head and backskins were pre-fixed in 4% PFA for 4 h at 4 degrees, followed by washes in PBS and incubation in 30% sucrose, before embedding in OCT. Sections were blocked for 1 h in gelatin block (5% normal donkey serum, 1% BSA, 2% fish gelatin, 0.3% Triton X-100 in PBS). Primary antibodies were diluted in blocking buffer and incubated at 4 °C overnight (O/N). MOMBasic kit (Vector Laboratories) was used for blocking when primary antibodies were generated from mouse. After washing with PBS, secondary antibodies, were added for 1 h at room temperature (RT). Slides were washed with PBS, counterstained with DAPI and mounted in Prolong Gold (Invitrogen). Images were acquired with an Axio Observer.Z1 epifluorescence microscope equipped with a Hamamatsu ORCA-ER camera (Hamamatsu Photonics), and with an ApoTome.2 (Carl Zeiss) slider that reduces the light scatter in the fluorescent samples, using 20 \times objective, controlled by Zen software (Carl Zeiss). Z stacks were projected and RGB images were assembled using ImageJ. Panels were labelled in Adobe Illustrator CS5.

Lentiviral expression constructs. Lentiviral super-enhancer reporters were generated by PCR amplification of selected enhancer regions from BAC clones, followed by insertion into KpnI and BsaBI restriction sites of the *Rbpj-EGFP* construct³⁷. To generate the *Sox9* expression construct, *Sox9*cDNA was PCR amplified, and inserted into the *LV-TRE-PGK-H2BmRFP1* construct¹⁸. The resulting *LV-TRE-mycSox9-PGK-H2BmRFP* was used for *in utero* injections.

Partial thickness wound (dermabrasion) and hair follicle stem cell transplantation. Animals were anaesthetized with ketamine/xylazine and administered buprenorphine analgesia. Skin was shaved and remaining hair cleared with hair removal cream. Skin was gently stretched between two fingers and epidermis removed using a small rotary drill (Dremel) with a polishing wheel attachment (model 520), to create a partial-thickness wound. Hair follicle stem cell transplantations were described previously²⁶.

Cell culture. Primary hair follicle stem cells were isolated from P52-60 K14-H2B-iRFP mice and plated onto mitomycin C-treated dermal fibroblasts in E-media supplemented with 15% (vol/vol) serum and 0.3 mM calcium³². For colony formation assays, equal numbers of *Sox9*-deficient live cells were plated. After 14 days in culture, cells were fixed and stained with 1% (wt/vol) Rhodamine B (Sigma). Colony diameter was measured from scanned images of plates using ImageJ and colony numbers were counted. For viral infections, hair follicle stem cells were spun with lentivirus for 30 min at 1,100g in the presence of polybrene (100 μ g ml⁻¹)²³. For *Sox9* overexpression studies, the *PGK-Sox9-IRES-H2BYFP* construct was transfected into cultured hair follicle stem cells or epidermal keratinocytes. 72 h later, YFP⁺ and YFP⁻ cells were purified by FACS. Luciferase assays were performed as described¹⁵.

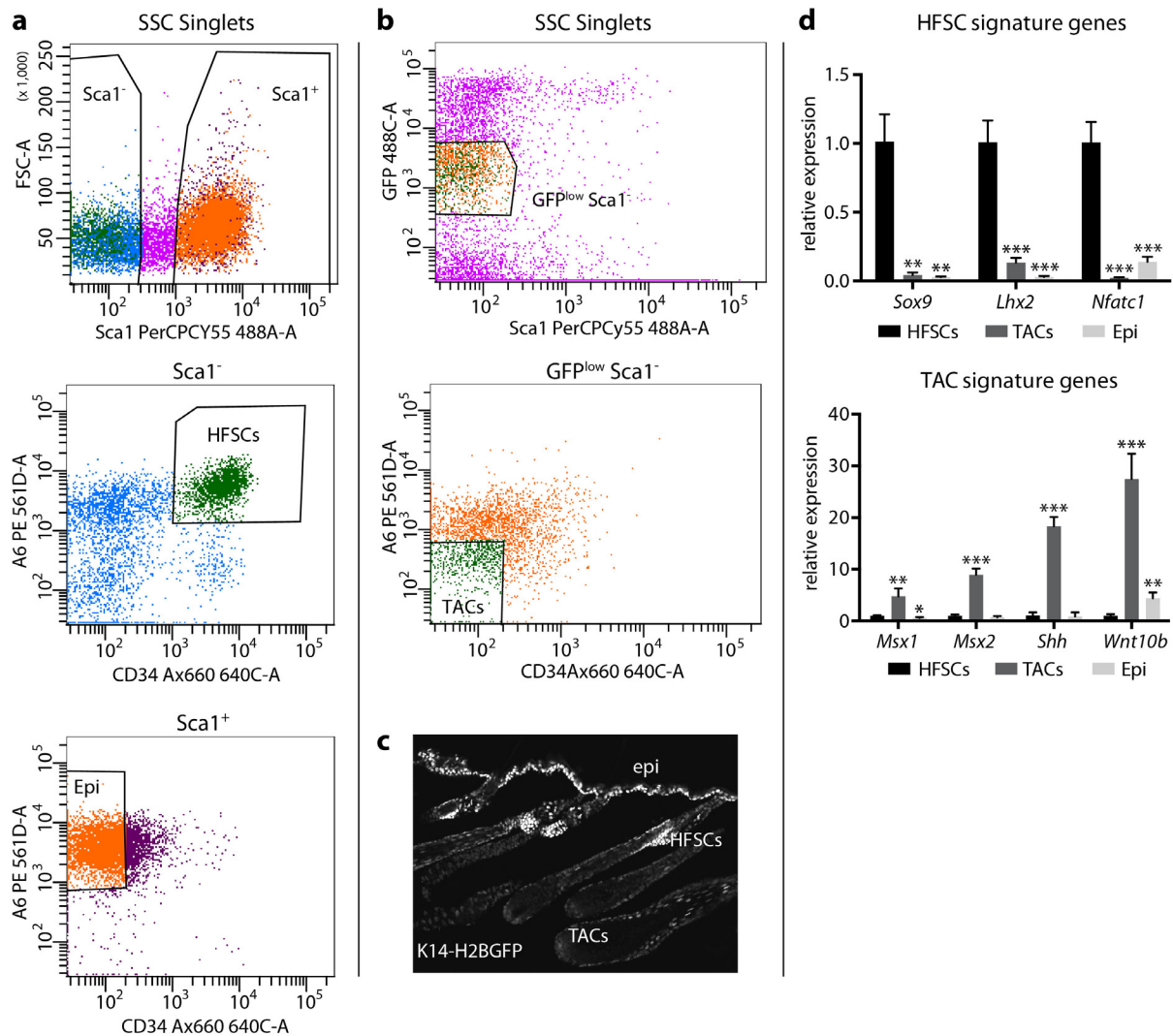
RNA extraction and qRT-PCR. FACS-isolated cells were sorted directly into TRIzol LS (Invitrogen). Total RNA was purified using the Direct-zol RNA MiniPrep kit (Zymo Research) per manufacturer's instructions. DNase treatment was performed to remove genomic DNA (RNase-free DNase Set, Qiagen). Equal amounts of RNA were reverse-transcribed using oligo-dT primers (Superscript III, Life Technologies). qRT-PCR was performed on an Applied Biosystems 7900HT Fast Real-Time PCR system. cDNAs were normalized to equal amounts using primers against Ppib2.

Statistics. For all measurements, 3 biological replicates and 2 or more technical replicates were used. Experiments were independently replicated twice, and representative data are shown. To determine the significance between two groups, comparisons were made using unpaired two-tailed Student's *t*-test in Prism6 (GraphPad software). For all statistical tests, the 0.05 level of confidence was accepted for statistical significance.

Sample size was predetermined based on the following considerations: at E9.5, surface ectoderm contained $\geq 120,000$ cells per embryo, each of which undergo 5–6 divisions until birth. We therefore estimate that the majority of hair follicles are derived from individual clones. Assuming 50% lentiviral infection efficiency at E9.5, we estimate $\geq 60,000$ independent lentiviral integration sites per animal. Being able to analyse ≥ 100 hair follicles per animal, we reasoned that transducing ≥ 2

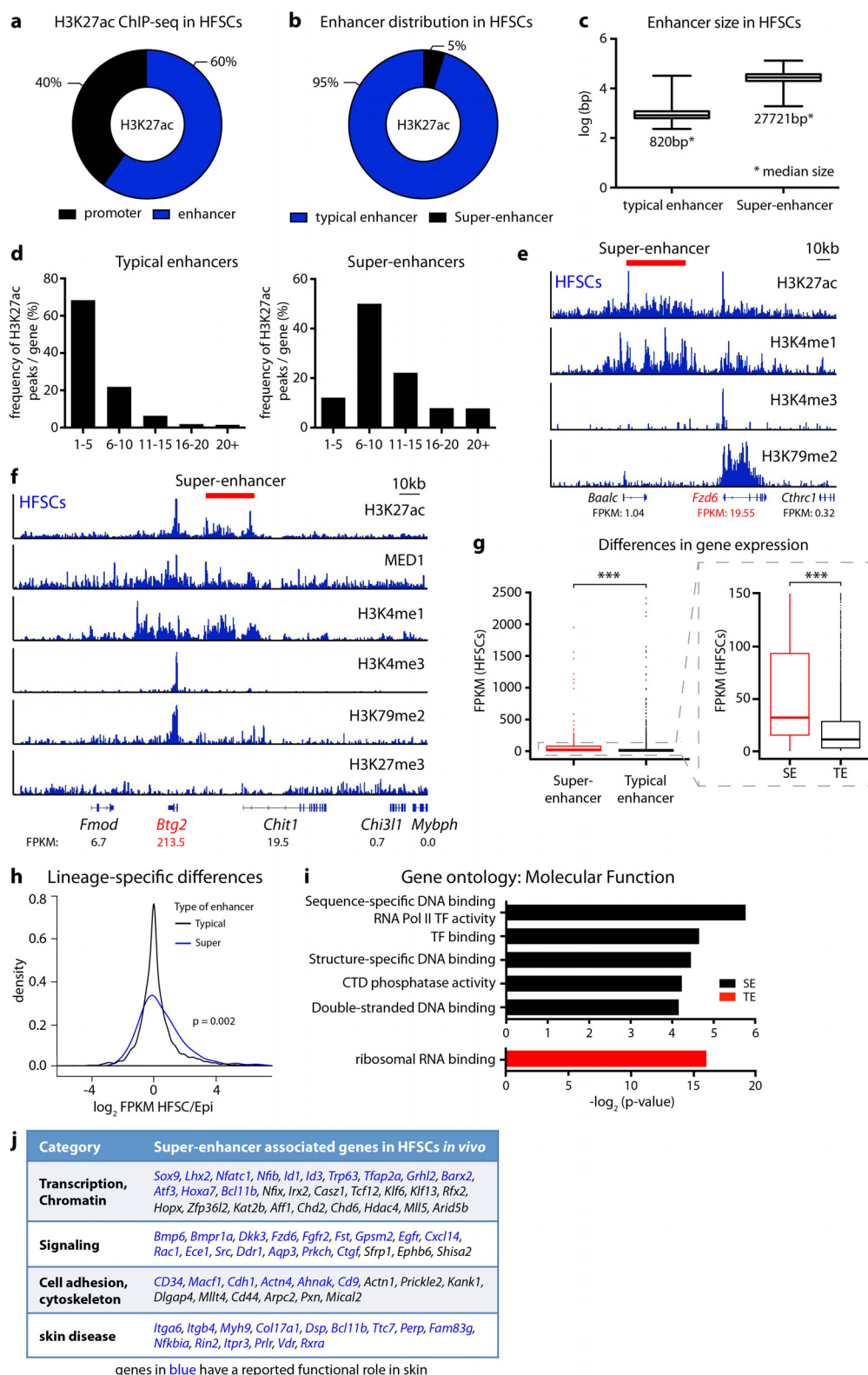
embryos from two separate litters would achieve the requisite coverage and control for any aberration due to a particular litter, independent of LV construct.

29. Tumber, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004).
30. Means, A. L., Xu, Y., Zhao, A., Ray, K. C. & Gu, G. A. CK19(CreERT) knockin mouse line allows for conditional DNA recombination in epithelial cells in multiple endodermal organs. *Genesis* **46**, 318–323 (2008).
31. Vasioukhin, V., Degenstein, L., Wise, B. & Fuchs, E. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl Acad. Sci. USA* **96**, 8551–8556 (1999).
32. Nowak, J. A. & Fuchs, E. Isolation and culture of epithelial stem cells. *Methods Mol. Biol.* **482**, 215–232 (2009).
33. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
34. Zhang, Y. *et al.* Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
35. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnol.* **28**, 495–501 (2010).
36. Ye, T. *et al.* seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* **39**, e35 (2011).
37. Williams, S. E., Beronja, S., Pasolli, H. A. & Fuchs, E. Asymmetric cell divisions promote Notch-dependent epidermal differentiation. *Nature* **470**, 353–358 (2011).



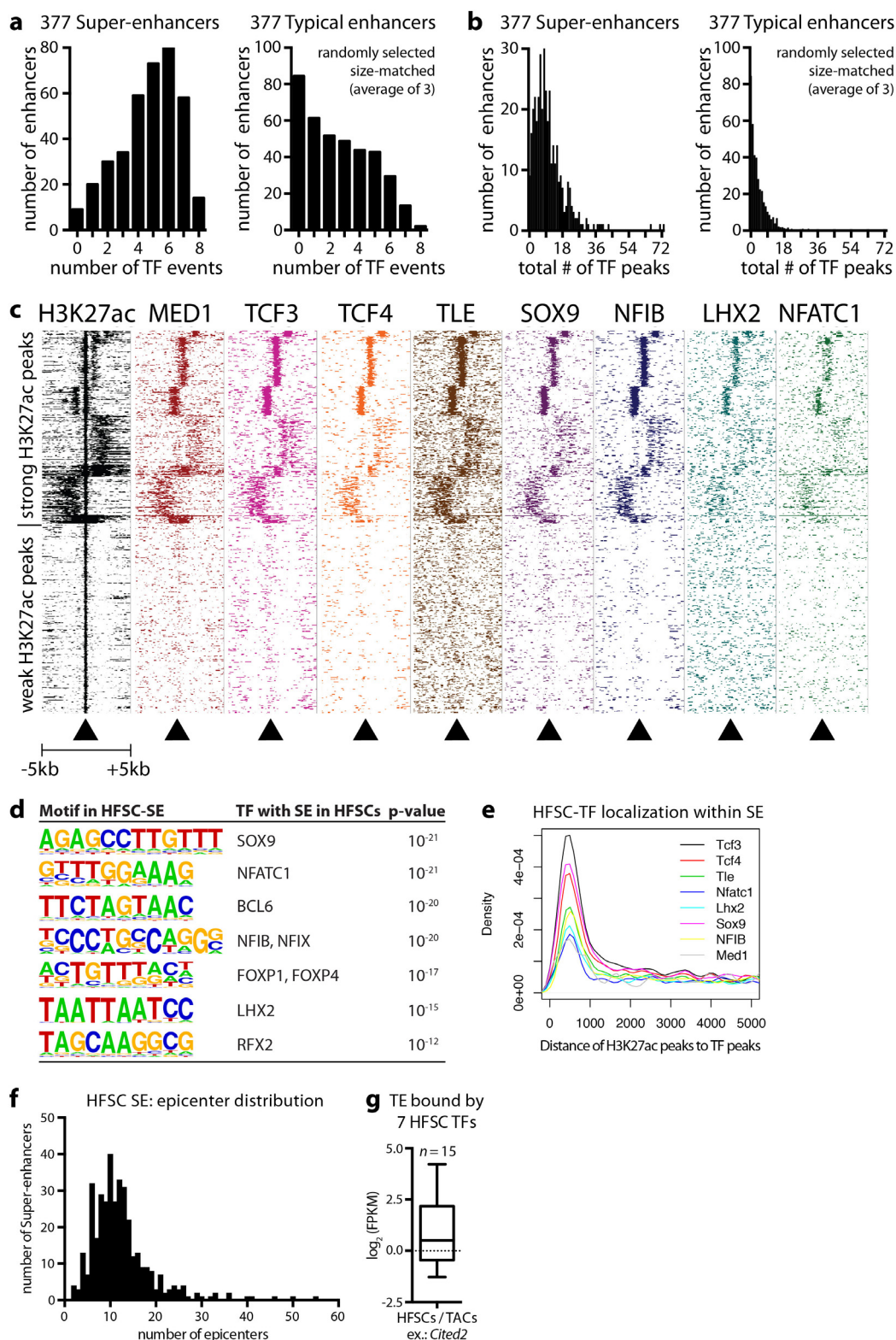
Extended Data Figure 1 | FACS purification strategy to isolate hair follicle stem cells and TACs. **a**, FACS purification of wild-type hair follicle stem cells for ChIP-seq according to established markers $\alpha 6^{hi}$ and CD34⁺²⁶. Sca1 is used to remove basal epidermal cells. **b**, FACS purification of TACs from Krt14-H2B-GFP mice²⁹. TACs are GFP^{lo}Sca1⁻ $\alpha 6^{lo/-}$ CD34⁻.

c, Epifluorescence of Krt14-driven H2B-GFP. Hair follicle stem cells and epidermal cells are GFP^{hi}, whereas TACs are GFP^{lo}. **d**, q-PCR to verify the FACS sorting strategy and measure enrichment of cell-type-specific marker genes. Mean and standard deviation are shown ($n = 3$). P values from t -test: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$, relative to hair follicle stem cells.



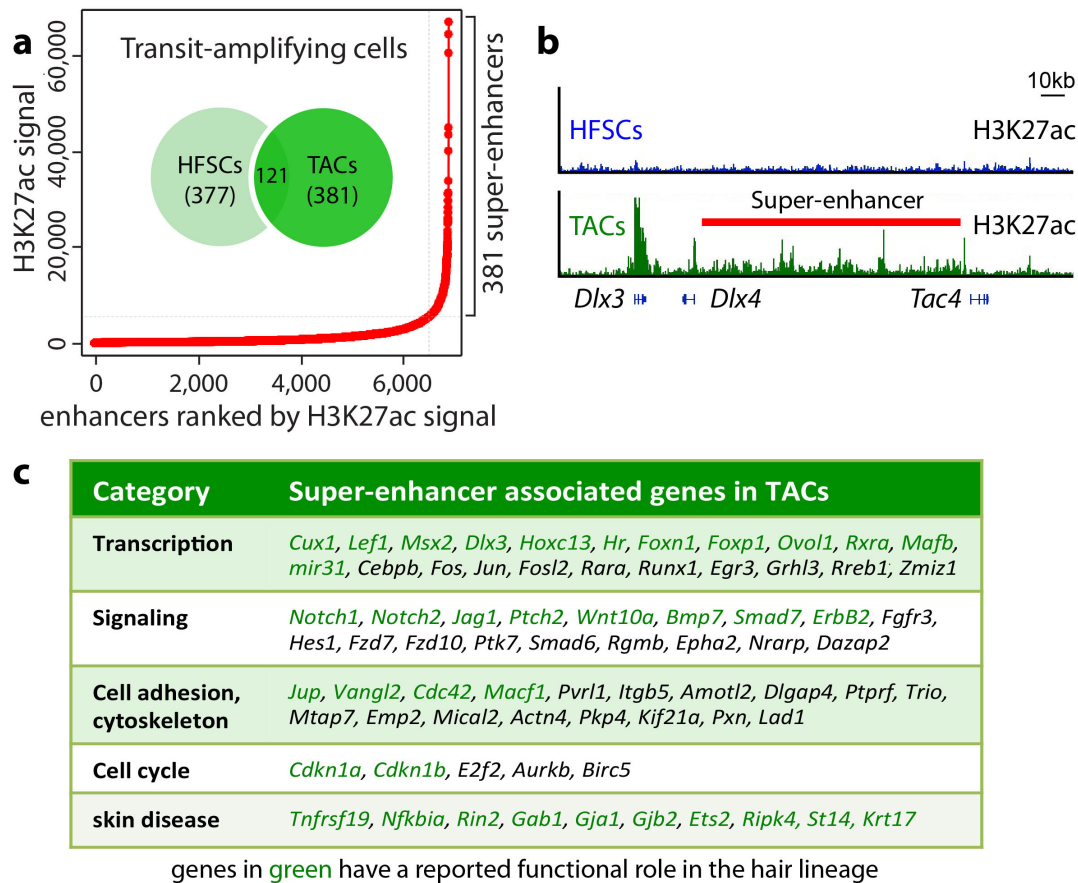
Extended Data Figure 2 | Enhancer distribution, size and gene assignment in hair follicle stem cells. **a**, Distribution of H3K27ac occupancy at promoter and enhancers in hair follicle stem cells. **b**, Distribution of typical-enhancers and super-enhancers in hair follicle stem cells. **c**, Enhancer size distribution in hair follicle stem cells. **d**, Number of individual H3K27ac peaks per gene. Super-enhancers are clusters of H3K27ac peaks and mainly consist of ≥ 5 peaks per gene. **e**, **f**, Enhancer-gene assignments, exemplified by hair follicle stem cell super-enhancers *Fzd6* and *Btg2*. FPKM, fragments per kilobase of transcript per million mapped reads (RNA-seq). **g**, Differential expression for

genes driven by hair follicle stem cell super-enhancers and typical-enhancers. *P* values from *t*-test: ****P* < 0.001. **h**, Density plot, contrasting expression levels of typical-enhancer versus super-enhancer associated hair follicle stem cell genes in hair follicle stem cells compared to epidermal progenitors. Note cell type-specific differences in expression for hair follicle stem cell genes controlled by super-enhancers but not typical-enhancers. **i**, Gene Ontology analysis of genes controlled by hair follicle stem cell enhancers. **j**, List of selected super-enhancer regulated hair follicle stem cell genes. SE, super-enhancer; TE, typical-enhancer.



Extended Data Figure 3 | Hair follicle stem cell TFs are enriched within super-enhancers and cluster in epicentres. **a, b**, Enrichment of hair follicle stem cell TFs within chromatin of super-enhancers, but not typical-enhancers. Comparisons were made with 377 randomly selected typical-enhancers and their flanking sequence extended 5' and 3' to match the average length of super-enhancers (average of 3 analyses is shown). Each 'TF event' (**a**) represents one hair follicle stem cell TF bound within a super-enhancer. 'TF peaks' (**b**) refers to the absolute amount of TFs occupying the super-enhancer. **c**, Heatmap showing ChIP-seq read densities (from -5 kb to +5 kb of peak centre) across H3K27ac peaks located in super-enhancers. Note that hair

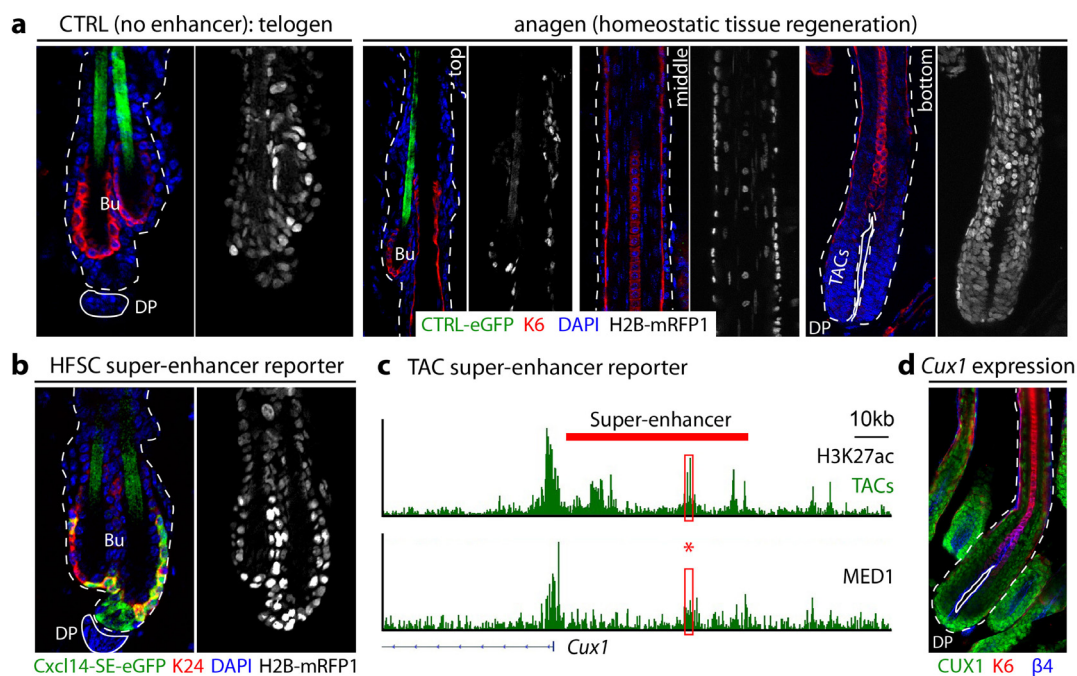
follicle stem cell TFs frequently bound densely together with strong H3K27ac peaks. **d**, Motif analysis of hair follicle stem cell super-enhancers for putative TF binding sites. **e**, Analysis of distance of H3K27ac peaks to their nearest transcription factor ChIP-seq peaks in hair follicle stem cells *in vivo* (distance of the two peak centres). Note that enrichment of TF binding occurs within 1-kb regions of H3K27ac peaks ('epicentres'). **f**, Frequency and distribution of hair follicle stem cell super-enhancer epicentres. **g**, Rare 'atypical' enhancers co-bound by 7 hair follicle stem cell TFs are more highly expressed in hair follicle stem cells versus committed progenitors.



Extended Data Figure 4 | Identification of super-enhancers in TACs.

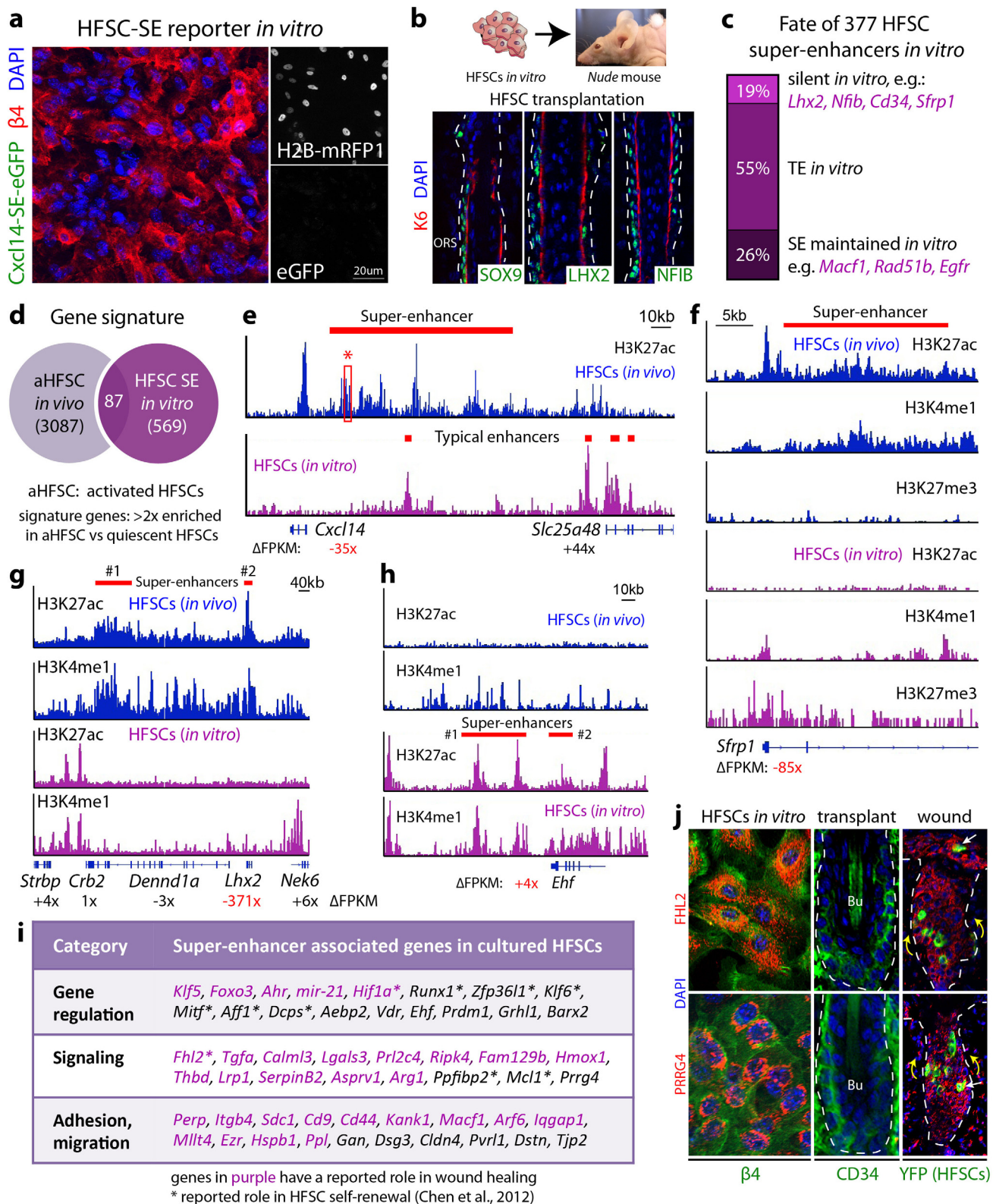
a, Distribution of H3K27ac ChIP-seq signal across all enhancers in transit-amplifying cells (TACs) reveals 381 super-enhancers of little overlap with hair follicle stem cell super-enhancers. **b**, Tracking the status of TAC super-enhancers in hair follicle stem cells indicates striking enhancer remodelling

upon lineage progression. Example shows the appearance of a *de novo* super-enhancer for the *Dlx3/4* locus as hair follicle stem cells commit to a TAC fate. **c**, Examples of super-enhancer associated genes in TACs. Genes in green have a reported function in hair follicles.



Extended Data Figure 5 | Super-enhancer reporters drive cell-type specific expression. **a**, The lentiviral control (CTRL) reporter construct (containing no enhancer) is silent throughout all stages of the hair cycle, despite efficient infection (as evidenced by H2B-mRFP1). **b**, Immunofluorescence showing that *Cxc14-eGFP* super-enhancer reporter activity co-localizes with Krt24⁺ hair follicle stem cells. DP, dermal papilla; Bu, Bulge. White dashed lines denote the

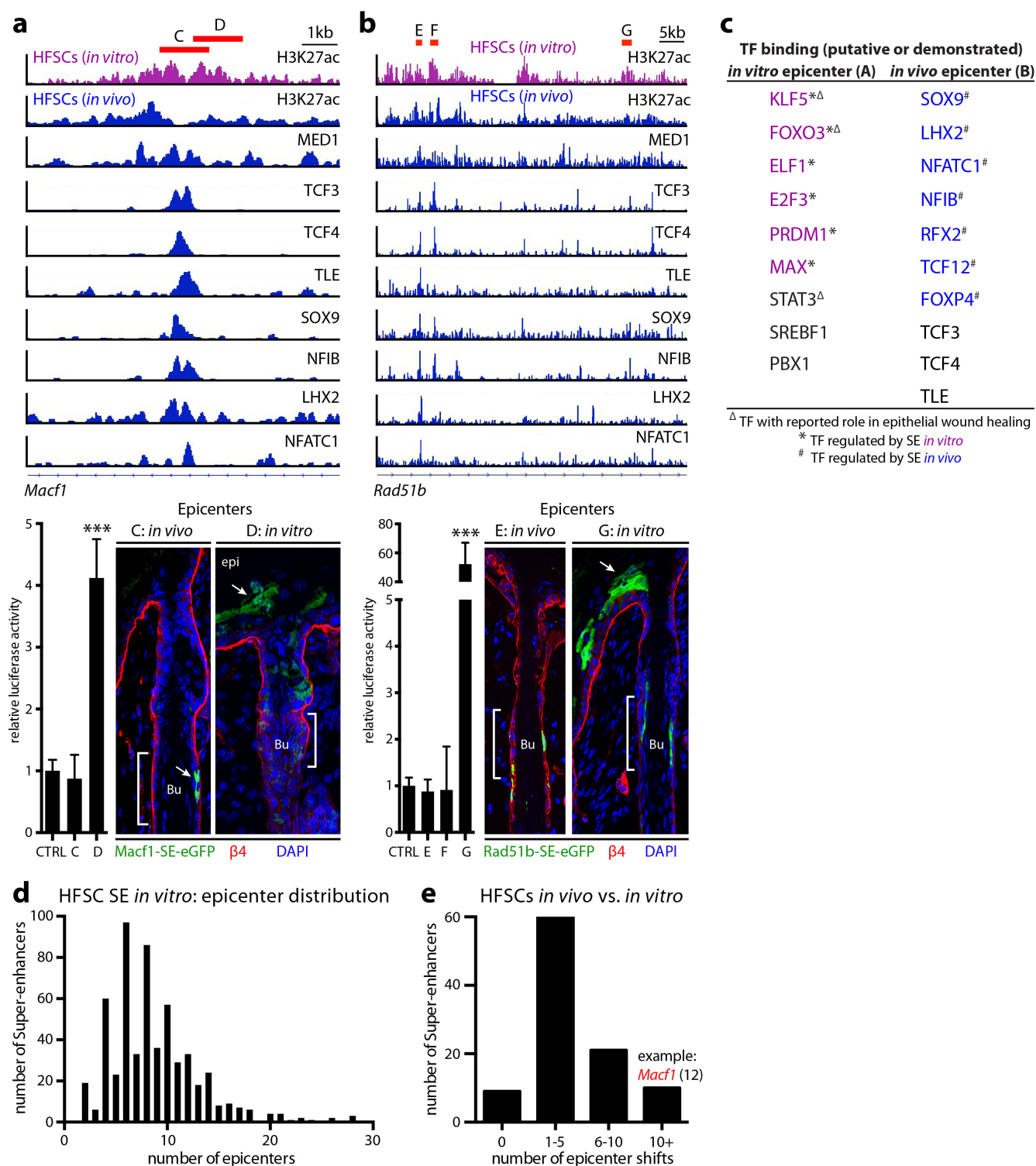
epidermal–dermal border; solid lines delineate the DP. **c**, H3K27ac and MED1 ChIP-seq occupancy at the *Cux1* locus in TACs. Red box shows the super-enhancer epicentre that was cloned for reporter assays. Note that epicentres bound by MED1 are sufficient to identify cell-stage specific loci, even without prior information about lineage-specific TFs. **d**, CUX1 expression pattern in hair follicles.



Extended Data Figure 6 | Hair follicle stem cells adapt to microenvironmental changes by reversible remodelling of super-enhancers.

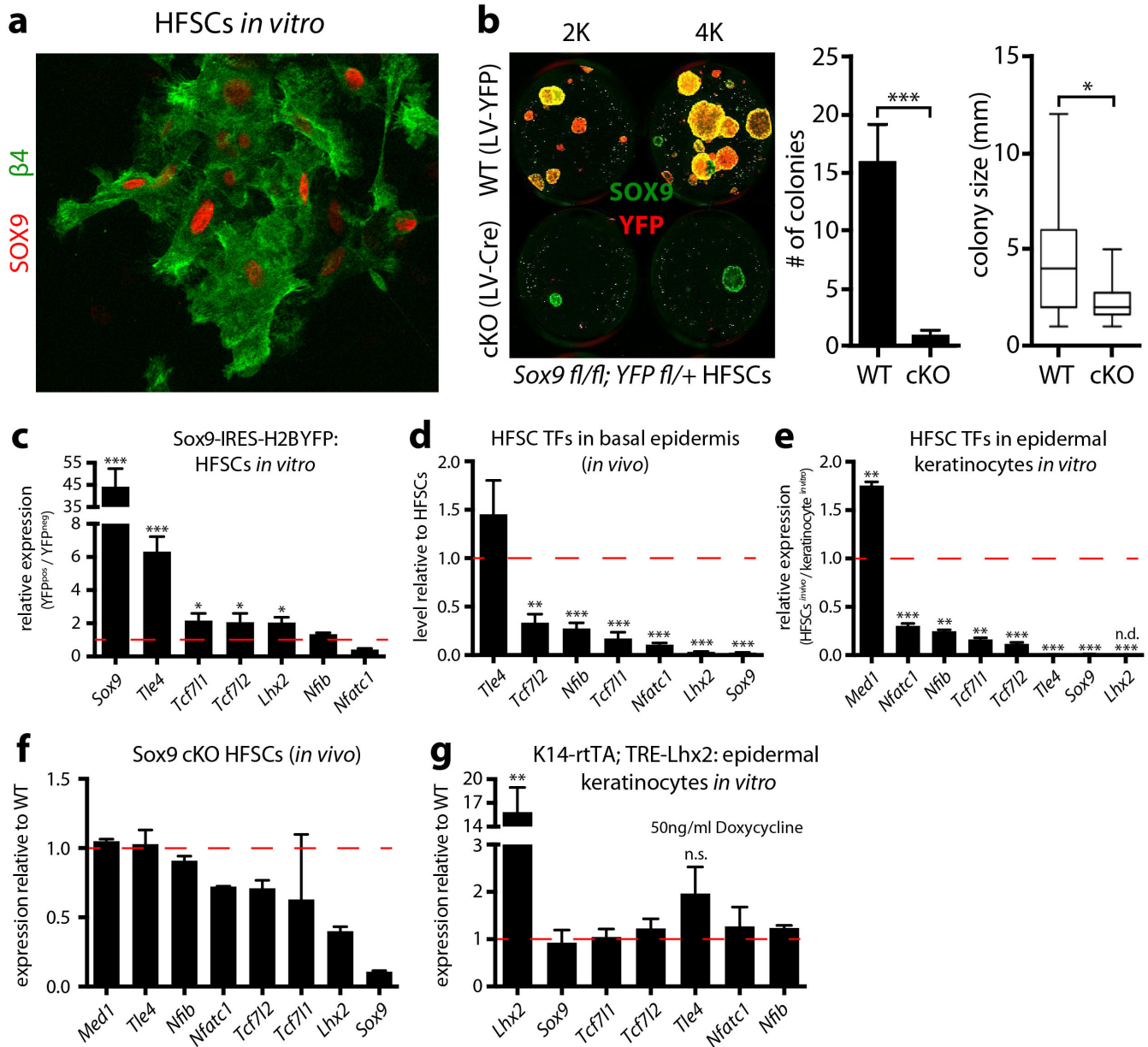
a, Absence of *Cxcl14*-SE-eGFP reporter activity in transduced cultured hair follicle stem cells. **b**, Transplanted cultured hair follicle stem cells establish *de novo* hair follicles and regain expression of hair follicle stem cell TFs. **c**, Note extensive hair follicle stem cell super-enhancer remodelling upon culture conditions. **d**, Hair follicle stem cells *in vitro* are molecularly distinct from activated hair follicle stem cells (aHFSC) *in vivo*. **e–h**, H3K27ac levels at the *Cxcl14*, *Sfrp1*, *Lhx2* and *Ehf* loci in hair follicle stem cells *in vivo* and *in vitro*.

Note the dynamic regulation of super-enhancers and the resulting changes in gene expression. **i**, Selected list of super-enhancer associated genes in hair follicle stem cells *in vitro*. **j**, Note hair follicle stem cell super-enhancer plasticity *in vitro* and during wound repair: *Fhl2* and *Prrg4* display super-enhancer-mediated activity *in vitro*. Upon transplantation, hair follicle stem cells silence *in vitro* induced genes concomitant with hair follicle regeneration. However, during wounding, hair follicle stem cells (lineage marked with K19-CreER/R26YFP) regain expression of *Fhl2* and *Prrg4*.



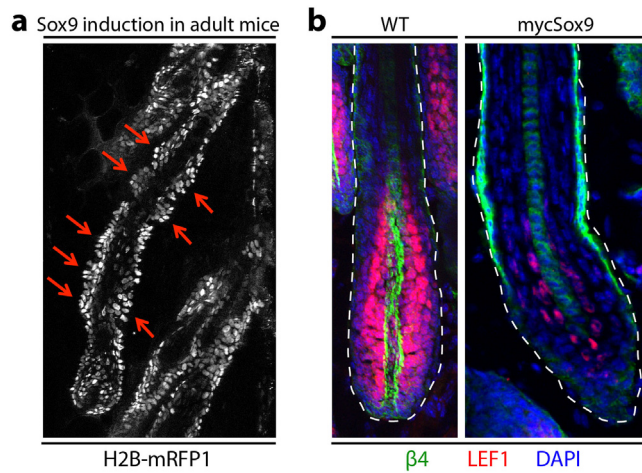
Extended Data Figure 7 | Hair follicle stem cells activate different epicentres within super-enhancers to sustain expression of critical genes in different microenvironments. **a**, **b**, H3K27ac and hair follicle stem cell TF ChIP-seq occupancies at the *Macf1* and *Rad51b* loci in hair follicle stem cells *in vivo* and *in vitro*. Regions C, E and F mark epicentres active *in vivo*, richly bound by hair follicle stem cell TFs; adjacent regions D and G are novel epicentres active *in vitro*. Relative luciferase activities were driven by the 1–1.5 kb encompassing these epicentres. Mean and standard deviation are shown ($n = 3$). P values from t -test: *** $P < 0.001$. Functional validation of epicentre shifts *in vivo*.

eGFP-reporter activity of *in vitro* epicentres is highly active in the epidermis, while physiological hair follicle stem cell epicentres are restricted to the hair follicle niche. **c**, Motif analysis of *Macf1* epicentres (regions A and B, Fig. 3e) for putative TF binding sites. **d**, Number and distribution of hair follicle stem cell super-enhancer epicentres *in vitro*. **e**, Frequency of epicentre shifts in hair follicle stem cell super-enhancers (*in vivo* versus *in vitro*). Note that corresponding to the loss of hair follicle stem cell TFs *in vitro*, many super-enhancers display epicentre shifts to maintain expression of critical genes (for example, *Macf1*) in different microenvironments.



Extended Data Figure 8 | Hair follicle stem cell TFs are reduced outside the niche but are sensitive to *Sox9* levels. **a**, SOX9 is expressed and displays nuclear localization in hair follicle stem cells *in vitro*. **b**, Colony formation assays on wild-type and *Sox9*-cKO hair follicle stem cells. *Sox9^{fl/fl}* Rosa26YFP^{fl/+} hair follicle stem cells were seeded at 2,000 and 4,000 and transduced with lentiviral-Cre to achieve *Sox9* ablation *in vitro*. All yellow and green colonies were not effectively targeted and are still SOX9⁺. All red colonies (SOX9-negative) aborted, as revealed by quantifications of colony numbers and sizes shown at right. **c**, *Sox9*-overexpression in cultured hair follicle stem

cells. SOX9 induces the expression of *Tle4*, *Tcf7l1*, *Tcf7l2* and *Lhx2*. **d**, **e**, Hair follicle stem cell TFs are expressed at substantially lower levels in basal epidermal progenitors *in vivo* or in cultured epidermal keratinocytes relative to hair follicle stem cells. **f**, Downregulation of hair follicle stem cell TFs in *Sox9*-cKO hair follicle stem cells *in vivo* before hair follicle stem cells are lost. **g**, Doxycycline-inducible overexpression of *Lhx2* in cultured epidermal keratinocytes does not induce hair follicle stem cell TFs. For **b–g**, mean and standard deviation are shown ($n = 3$). P values from t -test: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; n.d., not detected; n.s., not significant.



Extended Data Figure 9 | Sustained Sox9 expression in committed progenitors perturbs lineage progression. **a**, Sustained *Sox9* in adult mice (doxycycline for 3 weeks in adult mice, starting at P21) leads to *de novo* formation of minibulge-like structures along the ORS. **b**, Immunofluorescence showing that *Lef1* (normally H3K27me3 repressed in hair follicle stem cells, but H3K27ac super-enhancer induced in TACs) remains repressed in mycSOX9⁺ hair follicles.

Recursive splicing in long vertebrate genes

Christopher R. Sibley^{1,2*}, Warren Emmett^{3*}, Lorea Blazquez¹, Ana Faro⁴, Nejc Haberman¹, Michael Briese^{2,5}, Daniah Trabzuni^{1,6}, Mina Rytén^{1,7}, Michael E. Weale⁷, John Hardy¹, Miha Modic^{2,8}, Tomaž Curk⁹, Stephen W. Wilson⁴, Vincent Plagnol³ & Jernej Ule^{1,2}

It is generally believed that splicing removes introns as single units from precursor messenger RNA transcripts. However, some long *Drosophila melanogaster* introns contain a cryptic site, known as a recursive splice site (RS-site), that enables a multi-step process of intron removal termed recursive splicing^{1,2}. The extent to which recursive splicing occurs in other species and its mechanistic basis have not been examined. Here we identify highly conserved RS-sites in genes expressed in the mammalian brain that encode proteins functioning in neuronal development. Moreover, the RS-sites are found in some of the longest introns across vertebrates. We find that vertebrate recursive splicing requires initial definition of an 'RS-exon' that follows the RS-site. The RS-exon is then excluded from the dominant mRNA isoform owing to competition with a reconstituted 5' splice site formed at the RS-site after the first splicing step. Conversely, the RS-exon is included when preceded by cryptic promoters or exons that fail to reconstitute an efficient 5' splice site. Most RS-exons contain a premature stop codon such that their inclusion can decrease mRNA stability. Thus, by establishing a binary splicing switch, RS-sites demarcate different mRNA isoforms emerging from long genes by coupling cryptic elements with inclusion of RS-exons.

Recursive splicing has been validated within the long introns (>24 kilobases (kb)) of three *D. melanogaster* genes^{1,2}. The RS-sites in these introns contain a 3' splice site followed by a sequence that reconstitutes a 5' splice site after the first part of the intron is spliced, thereby allowing subsequent splicing of the second part of the intron (Fig. 1a). While one mammalian sequence located at the start of an alternative exon was proposed to function as an RS-site when pre-spliced to an upstream exon in a splicing reporter³, recursive splicing has not been observed in endogenous vertebrate genes. This is despite >8,000 human protein-coding genes containing introns >24 kb, and many vertebrate genes containing motifs similar to the *D. melanogaster* RS-sites⁴.

Long genes exhibit increased expression in the nervous system, as evident by analysis of human tissues or differentiating cells⁵ (Fig. 1b and Extended Data Fig. 1b–d), and are enriched in Gene Ontology (GO) terms associated with the nervous system (Extended Data Fig. 1a). We therefore produced 1.5 billion paired-end total RNA sequencing (RNA-seq) reads from four post-mortem brains to search for new splicing events in human long genes. Notably, RNA abundance decreases linearly from the 5' to 3' end of long introns to create 'saw-tooth' patterns in total RNA-seq data⁶, and these can be used to infer locations of major splicing events (Fig. 1c, d and Extended Data Figs 2a and 3). We also performed crosslinking and immunoprecipitation (iCLIP) of the RNA-binding protein fused in sarcoma (*FUS*) in human brain. *FUS* binds across entire pre-mRNAs with limited sequence specificity⁷, permitting an independent examination of the saw-tooth patterns (Fig. 1d and Extended Data Fig. 3a–g).

Cryptic splice sites can be identified from novel splice-junction reads in RNA-seq data (Extended Data Fig. 2c–e). We proposed that if some of these were major splicing events, they should cause significant deviations from the expected linear decrease of reads across long introns (Fig. 1c, d). Analysis of our RNA-seq data identified 40,163 unique, unannotated cryptic splice sites in introns >1 kb that contained either 5' or 3' splice site motifs, 419 of which conformed to the RS-site motif (Supplementary Table 1, worksheets 1 and 2). We evaluated deviations from the expected saw-tooth pattern by establishing an analysis that computed the fit of linear regression slopes of each intron as a single unit or as two units separated at newly detected intronic junctions (Fig. 1c–e and Extended Data Figs 2a, b and 3). Since intron size is a crucial determinant of our ability to detect unexpected saw-tooth patterns reliably, we restricted our analysis to genes with at least one intron >150 kb. This identified 19 unique cryptic splice sites in the long introns of 14 genes that significantly improved the goodness-of-fit of the regression model in both RNA-seq and *FUS* iCLIP data sets. Of these, 9 had the RS-site motif while the remainder had a 3' splice site motif ($P < 0.01$ in both data sets, Fig. 1d–f and Supplementary Table 1, worksheet 3). The genes containing these 9 RS-sites mostly function in cell adhesion and axon guidance and are linked to neurodevelopmental disorders (Supplementary Table 2).

The 9 RS-sites occurred at transition points of intronic linear regression slopes in all four individuals and all brain regions profiled (Fig. 1d and Extended Data Figs 3 and 4). Reverse transcription PCR (RT-PCR) from a separate human brain validated splice products for 8 RS-sites which were detectable at identical PCR cycle number as the mature mRNA, suggesting equal abundance, while no PCR products were observed when reverse primers were shifted upstream of RS-sites (Fig. 2a and Extended Data Fig. 5a–g).

Notably, an alternative 5' splice site is present downstream of each RS-site that could lead to inclusion of alternative exons (hereafter termed RS-exons, Fig. 2b). However, RS-exons were not detectable in mRNA transcripts at comparable PCR cycle numbers used to detect RS-site junctions (Fig. 2a and Extended Data Fig. 5a–g), arguing that RS-sites are being used for recursive splicing rather than for RS-exon inclusion. Despite RS-exon skipping, mammalian conservation of both the RS-sites and alternative 5' splice sites following the RS-exons is comparable to that of canonical 5' and 3' splice sites (Fig. 2c, d and Extended Data Fig. 5i). Indeed, mouse *Fus* iCLIP regression patterns directly match conserved RS-sites⁷ (Extended Data Fig. 6a–h).

Splicing of most vertebrate exons requires exon definition⁸, in which both splice sites flanking an exon are recognized in unison via interactions between U2AF proteins, Ser/Arg-rich (SR) proteins and small nuclear ribonucleoproteins (snRNPs)⁹ (Supplementary Information). We speculated that RS-exons co-evolved with RS-sites to enable exon definition (Fig. 2e). Accordingly, we masked the 5' splice site following the *CADM1* and *ANK3* RS-exons in SH-SY5Y neuroblastoma cells using an antisense oligonucleotide (AON-A1; Fig. 2e). This markedly

¹Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK. ²MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.

³University College London Genetics Institute, Gower Street, London WC1E 6BT, UK. ⁴Department of Cell and Developmental Biology, University College London, Gower Street, London WC1E 6BT, UK.

⁵Institute for Clinical Neurobiology, University of Würzburg, Versbacherstr. 5, 97078, Würzburg, Germany. ⁶Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia. ⁷Department of Medical & Molecular Genetics, King's College London, Guy's Hospital, London SE1 9RT, UK. ⁸Institute of Stem Cell Research, German Research Center for Environmental Health, Helmholtz Center Munich, 85764 Neuherberg, Germany. ⁹Faculty of Computer and Information Science, University of Ljubljana, 1000 Ljubljana, Slovenia.

*These authors contributed equally to this work.

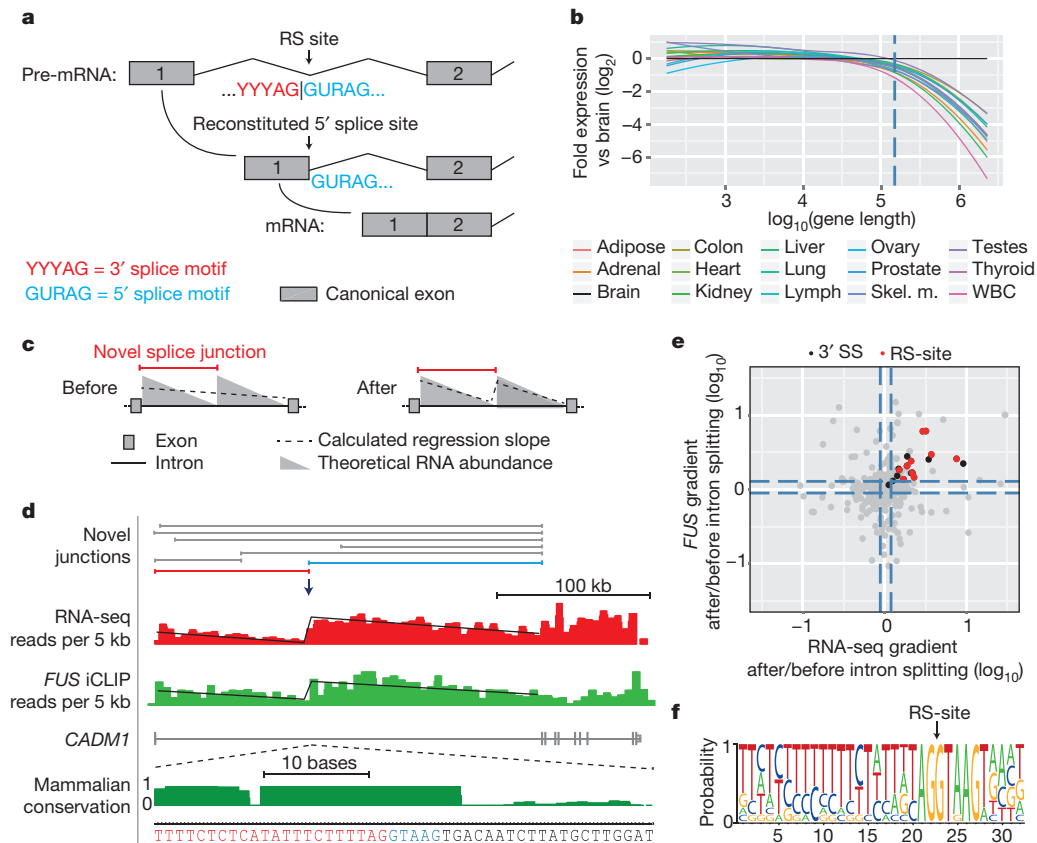


Figure 1 | Detection of recursive splice sites within long genes expressed in the human brain. **a**, Schematic of the *D. melanogaster* recursive splicing mechanism. **b**, The \log_2 -fold gene expression ratios following differential expression sequencing (DESeq)¹⁹ analysis of all human protein-coding genes between the brain and all other tissues. Data are represented as Loess smoothing curves after defining genes by their maximum length in kilobases. Dashed vertical line indicates 150 kb. RNA-seq data were obtained from the Illumina Human Body Map 2.0 total RNA-seq library (GEO accession GSE30611). Skel. m., skeletal muscle; WBC, white blood cell. **c**, Schematic of the theoretical RNA abundance across long introns demonstrating linear regression analysis performed on introns before/after novel junction consideration. **d**, All novel junctions identified within *CADM1* by RNA-seq data are shown on top of experimentally derived RNA-seq (red) and FUS iCLIP

(green) read densities, both grouped in 5-kb windows. The displayed linear regression line was determined after the intron was split at the red novel junction. This split significantly improved the regression in both RNA-seq and FUS iCLIP ($P < 0.01$ in both, F -test). Blue novel junction contacts the RS-exon. Phylo-P sequence conservation scores are shown around the *CADM1* RS-site across 46 mammalian species. **e**, Ratio of after: before gradients at long gene novel junctions in RNA-seq (x axis) and FUS iCLIP (y axis) data sets. Black and red dots represent junctions that significantly improve the regression gradient and goodness-of-fit, whereas grey dots show no improvement. Black dots are junctions contacting the sequence of 3' splice sites (SS), whereas red dots contact the sequence of RS-sites. Dashed lines mark top and bottom quartile ratios for each data set. **f**, WebLogo of RS-sites identified by red junctions from **e**.

reduced RS-site usage in both genes (Fig. 2f). We subsequently replicated this observation *in vivo* at the conserved RS-site/RS-exon of the zebrafish *cadm2a* gene (Fig. 2g and Extended Data Fig. 5h). The reduced RS-site usage also led to a ~ 6 -fold increase in abundance of the intronic region upstream of both human RS-sites, indicating a change in the saw-tooth pattern consistent with splicing of the intron as a whole (Fig. 2h). Interestingly, the reduced RS-site usage caused a ~ 2 -fold reduction in zebrafish *cadm2a* total mRNA (Fig. 2i), an effect not seen for the human *CADM1* and *ANK3* genes (Fig. 2j and Supplementary Information). Despite RS-exons usually being skipped, our findings demonstrate that RS-exon definition is crucial for the initial step of vertebrate recursive splicing (Fig. 2e).

Since recursive splicing requires initial definition of an RS-exon, we questioned whether some annotated alternative exons might function as RS-exons. We found 99 candidate annotated RS-exons with RS-site sequences located precisely at their starts (Extended Data Fig. 7a). Splice-junction reads from brain RNA-seq data were present at the start of 16 of these exons despite evidence for exon skipping. These included exons in the *CADM2* and *NTM* genes that significantly improved the goodness-of-fit of linear regression in RNA-seq and iCLIP data sets across their >150 kb introns (Supplementary Table 1,

worksheet 4 and 5). We confirmed RS-site mediated exon-skipping in both genes by RT-PCR (Extended Data Fig. 7b, f). Thus, the first intron in *CADM2* gene contains two RS-sites; the first followed by an unannotated RS-exon, and the second by an annotated RS-exon.

To validate the exon definition mechanism further, we established a splicing reporter containing the second *CADM2* RS-site, the annotated RS-exon and its 5' splice site, and the surrounding constitutive exons, each flanked by their nearest ~ 100 nucleotides of *CADM2* intronic sequence (P1; Fig. 3a). Despite the >500 -kb long intron being reduced to ~ 0.5 kb, the reporter replicated the findings of endogenous genes; 79% of mRNA isoforms skipped the RS-exon while RS-site usage was readily detected (Fig. 3b and Extended Data Fig. 8a). As expected given the need for exon definition to recognize RS-sites, mutating the 5' splice site following the RS-exon greatly reduced RS-site usage, and the intron remained a single unit in most splicing intermediates (P1-m1; Fig. 3a, b and Extended Data Fig. 8a).

Next, to examine why RS-exons are excluded from the mRNA, we mutated the RS-site of the *CADM2* reporter to prevent formation of the reconstituted 5' splice site after the first splicing event (Fig. 3a). Notably, this resulted in complete inclusion of the RS-exon, suggesting competition exists between the two 5' splice sites at either end of the

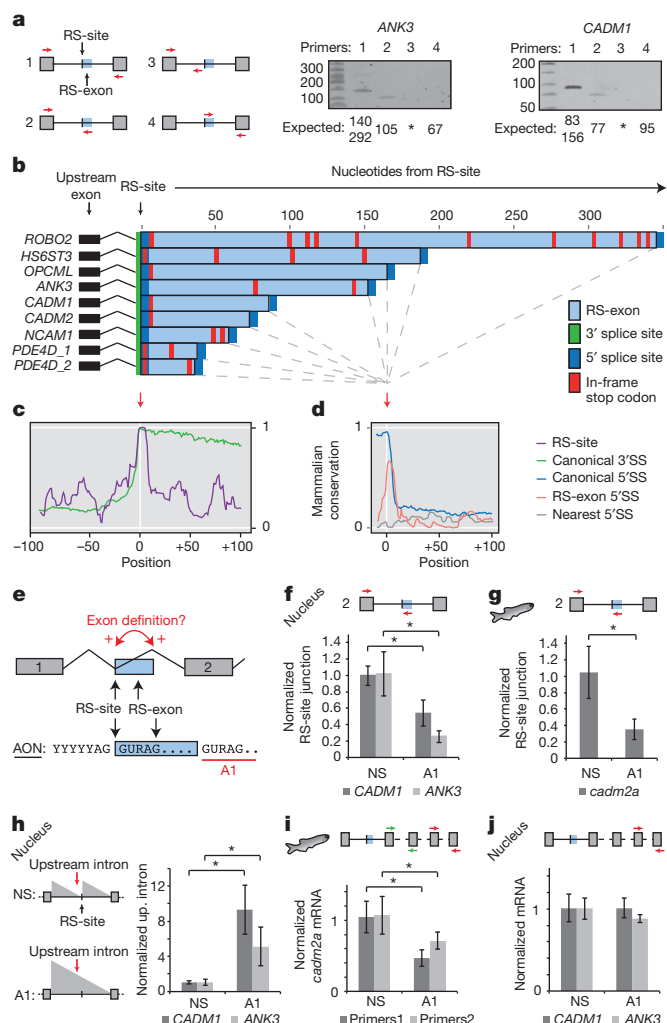


Figure 2 | Recursive splicing requires initial definition of RS-exons. **a**, RT-PCR validation of recursive splicing in *ANK3* and *CADM1* genes. The length of expected products in nucleotides is marked below the gels. No products are expected in the lanes marked by the asterisk. **b**, Consensus splice site location and in-frame termination codons at RS-exons in indicated human genes. **c**, **d**, Phylo-P conservation scores aligned at RS-sites (**c**) and 5' splice sites (**d**) of RS-exons. Conservation at the two nearest cryptic 5' splice sites following RS-exons (nearest 5' splice site) and the canonical 5' and 3' splice sites in the same genes are also shown. **e**, Schematic of the exon definition model and AON-A1 design strategy. **f**, **g**, Quantitative RT-PCR (qRT-PCR) analysis of RS-site junctions in human *CADM1* and *ANK3* genes ($n = 4$ for non-specific AON (NS), $n = 5$ for AON-A1, 2 separate experiments) (**f**) or zebrafish *cadm2a* gene after treatment with AON-A1 ($n = 7$, 3 separate experiments) (**g**). **h**, qRT-PCR analysis of intronic RNA upstream of RS-sites in *CADM1* and *ANK3* genes after treatment with AON-A1. Location of primer pair is indicated by red arrow in schematic, and expected changes in intronic abundance indicated by grey triangles ($n = 4$ for NS, $n = 5$ for AON-A1, 2 separate experiments). **i**, qRT-PCR analysis of zebrafish *cadm2a* mRNA using two separate primer sets targeting constitutive exons after *in vivo* injection of AON-A1 ($n = 7$, 3 separate experiments). **j**, qRT-PCR analysis of human *CADM1* and *ANK3* mRNAs after 48 h treatment with AON-A1. mRNA for both genes was assessed in nuclear fractions ($n = 4$ for NS, $n = 5$ for AON-A1, 2 separate experiments). $*P < 0.05$ (two-tailed student *t*-test). Data are mean \pm s.d. Unless indicated otherwise, primers are indicated by coloured arrows within schematics. Replicate data are shown in the source data.

RS-exon (P1-m2; Fig. 3a, b). To compare with endogenous genes, we designed another antisense oligonucleotide, AON-A2, to mask the section of RS-sites that contributes to the reconstituted 5' splice site in the human *CADM1*, *ANK3* or zebrafish *cadm2a* genes (AON-A2; Fig. 3a). Agreeing with the splicing reporter, AON-A2 markedly

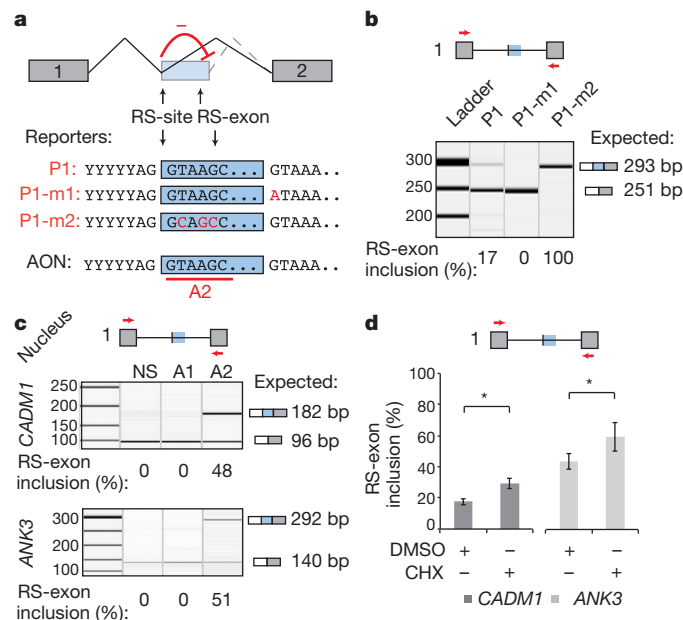


Figure 3 | The reconstituted 5' splice site is required for RS-exon skipping. **a**, Schematic of splice site competition model, and the design strategy for the *CADM2* splicing reporter P1 variants and AON-A2 experiments. **b**, **c**, Qiaxcel analysis of indicated *CADM2* splicing reporter products after transfection in SH-SY5Y cells ($n = 3$ –5, 2 separate experiments) (**b**), or human *CADM1* and *ANK3* genes after 48-h treatment with AON-A2 ($n = 4$ for *CADM1*, $n = 5$ for *ANK3*, 2 separate experiments) (**c**). **d**, Quantification of *CADM1* and *ANK3* RS-exon inclusion after treatment with AON-A2 then dimethylsulphoxide (DMSO) or cycloheximide (CHX) in SH-SY5Y cells ($n = 4$, 2 separate experiments). $*P < 0.05$ (two-tailed Student's *t*-test). Data are mean \pm s.d. Primers used are indicated by red arrows in schematics. Replicate data are shown in the source data.

increased RS-exon inclusion in all human and zebrafish experiments (Fig. 3c and Extended Data Fig. 8b). Collectively, this demonstrates that the RS-exon is skipped owing to a splice site competition that leads to use of the reconstituted 5' splice site instead of the 5' splice site of the RS-exon (Fig. 3a, and Supplementary Information).

We noticed that RS-exons typically contain one or more in-frame stop codons (Fig. 2b and Extended Data Fig. 5i), inclusion of which should prevent translation of full-length protein and target transcripts with preceding start codons to nonsense-mediated decay¹⁰. We induced inclusion of the RS-exons in *CADM1* and *ANK3* by masking the 5' splice site of their RS-sites with AON-A2, and then inhibited nonsense-mediated decay by blocking translation with cycloheximide. This increased the proportion of isoforms containing the RS-exon (Fig. 3d), confirming that RS-exon inclusion can target transcripts for nonsense-mediated decay and thus has the potential to regulate transcript stability (Supplementary Information).

Having identified the mechanisms underlying vertebrate recursive splicing, we next explored the functions of RS-sites. Although *D. melanogaster* RS-sites have been proposed to maintain splicing integrity of long introns⁴, the assayed human and zebrafish long introns remained accurately spliced after recursive splicing inhibition with AON-A1 (Extended Data Fig. 8c). We therefore explored an additional hypothesis that RS-sites regulate inclusion of RS-exons under specific contexts. We identified minor isoforms in the *CADM2* and *NTM* genes that use a different promoter, and were therefore not detected by our initial RT-PCR reactions. Their detection required 10 more amplification cycles compared to the dominant isoform, confirming that they are minor isoforms (Extended Data Fig. 7c, d, g). Surprisingly, RS-exons are completely included in these minor isoforms that have an alternative exon or promoter preceding the RS-site (Fig. 4a–c and

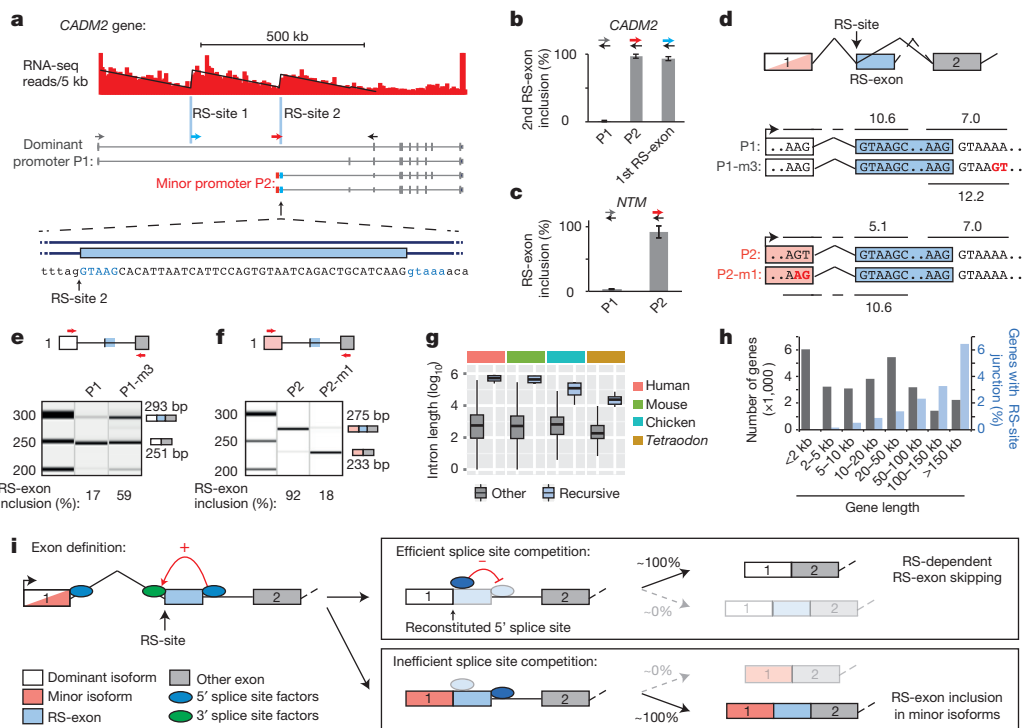


Figure 4 | Splice site competition allows a binary splicing switch for RS-exons. **a**, RNA-seq read density patterns in the CADM2 gene shown in 5-kb windows, with linear regression performed after the first intron is split at the two RS-sites indicated with blue vertical lines. Isoforms expressed from the dominant and minor promoters in human frontal cortex tissue are shown, and primer locations used for **b** indicated by coloured arrows. Grey forward primer is located in the first exon of dominant isoform, blue forward primer is located in the first RS-exon, red forward primer is located in the first exon of alternative isoform (P2). Zoomed area represents the sequence at the start of the second RS-exon. **b, c**, RT-PCR analysis of RS-exon inclusion in indicated CADM2 isoforms (**b**) or indicated NTM isoforms (**c**) ($n = 4$ and $n = 3$ respectively; Extended Data Fig. 7). Values are mean \pm s.d. **d**, Schematic of CADM2 splicing

Extended Data Fig. 7c–g). Similarly, the RS-exon is also detected in expressed sequence tags of minor *OPCML* isoforms that contain alternative exons preceding the RS-site (Extended Data Fig. 9a). A related splicing mechanism that coordinates alternative promoters with downstream alternative splicing has been observed in the human *EPB41* and *EPB41L3* genes, although this involves a reconstituted 3' splice site to make it distinct from recursive splicing¹¹.

To understand how preceding exons can dictate inclusion of RS-exons in a binary manner, we compared the computationally predicted strengths of the three relevant 5' splice sites in CADM2 (ref. 12); the 5' splice sites reconstituted from the RS-site after its splicing to the preceding exon of either the dominant or minor isoforms, and the 5' splice site of the RS-exon (Fig. 4d). We used the last three nucleotides of the preceding exon and the six nucleotides from the RS-site to calculate the scores of the reconstituted 5' splice sites¹². We found that the reconstituted 5' splice site had a high score when the first exon is derived from the dominant promoter (10.6), a low score when derived from the minor promoter (5.1), while the 5' splice site of the RS-exon had an intermediate score (7.0). This indicates that the strength of the reconstituted 5' splice site dictates whether the RS-exon is included or skipped. Indeed, 5' splice sites reconstituted from the preceding exon of the dominant isoform in all 9 high-confidence RS-sites had equal or higher splice site scores than the 5' splice sites of their corresponding RS-exons, in agreement with observed RS-exon skipping (Extended Data Fig. 8d and Supplementary Table 3, worksheet 1).

reporter variants P1 and P1-m3, based on the dominant CADM2 isoform (white), and P2 and P2-m1, based on the minor CADM2 isoform (red). Splice site scores for reconstituted and RS-exon 5' splice sites are indicated. **e, f**, Qiaxcel analysis of indicated CADM2 splicing reporter products after transfection in SH-SY5Y cells ($n = 3$ or $n = 4$, 2 separate experiments). The expected size of PCR products is shown next to each electropherogram. **g**, Lengths of the 9 introns containing high-confidence RS-sites compared to other vertebrate introns. **h**, Histogram of human gene lengths plotted alongside the percentage of genes with RS-site-containing novel junctions. **i**, Schematic representation of the mechanism of recursive splicing and the binary splicing switch as described in main text. For relevant panels, replicate data are shown in the source data.

To evaluate experimentally, we mutated the 5' splice site of the CADM2 RS-exon in our splicing reporter such that its score was higher (12.2) than the reconstituted 5' splice site of the dominant isoform (10.6, P1-m3; Fig. 4d). This mutation favoured RS-exon inclusion (Fig. 4e). We then replaced the preceding exon of the dominant isoform with the one from the minor isoform. This led to complete inclusion of the RS-exon, re-capitulating behaviour of the endogenous gene (P2; Fig. 4d, f). Finally, swapping the last three nucleotides of the preceding exon in the minor isoform to the sequence of dominant isoform led to RS-exon skipping, consistent with the higher score of the reconstituted 5' splice site (10.6, P2-m1; Fig. 4d, f). These results reveal that the binary splicing switch is a consequence of the relative strengths of competing 5' splice sites present after the RS-exon is spliced to the preceding exon.

Introns containing the high-confidence RS-sites are among the longest introns in all vertebrate species (Fig. 4g and Extended Data Fig. 9b). This includes *Tetraodon nigroviridis*, which has the shortest known vertebrate genome and otherwise contains very short introns¹³. Furthermore, 8 out of 9 of our high-confidence RS-sites are located in the long first intron of the gene. We confirmed that long introns generally have an increased incidence of cryptic exons and noisy splicing^{14,15} by observing an increased incidence of cryptic junctions in our RNA-seq data in long first introns (Extended Data Fig. 9c and Supplementary Table 3, worksheet 2 and 3). Because most of the 435 putative RS-sites identified in our study are present in the longest human genes (419 intronic loci, 16 annotated RS-exons; Fig. 4h), RS-

sites are thus well positioned to couple inclusion of cryptic exons with RS-exons. As most RS-exons contain a premature stop codon, this may also allow quality control of the novel mRNA isoforms (Supplementary Information).

In summary, recursive splicing of long vertebrate genes involves two steps (Fig. 4i). First, the RS-exon is defined, which requires its own 5' splice site. After splicing of the RS-exon to the preceding exon, a new 5' splice site is reconstituted from the RS-site that competes with the 5' splice site of the RS-exon. The strength of the reconstituted 5' splice site determines whether the RS-exon is skipped via recursive splicing or included. Notably, the upstream exons of dominant isoforms reconstitute a strong 5' splice site that leads to recursive splicing, whereas other alternative exons, which commonly emerge in long introns to produce minor isoforms, generally end in sequences that lead to RS-exon inclusion. In light of studies linking aberrant expression of long genes to neurological diseases^{16–18}, mutations or deletions around RS-sites may also contribute to human genetic diseases.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 May 2014; accepted 9 April 2015.

Published online 13 May 2015.

1. Burnette, J. M., Miyamoto-Sato, E., Schaub, M. A., Conklin, J. & Lopez, A. J. Subdivision of large introns in *Drosophila* by recursive splicing at nonexonic elements. *Genetics* **170**, 661–674 (2005).
2. Hatton, A. R., Subramaniam, V. & Lopez, A. J. Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon–exon junctions. *Mol. Cell* **2**, 787–796 (1998).
3. Grellscheid, S. N. & Smith, C. W. An apparent pseudo-exon acts both as an alternative exon that leads to nonsense-mediated decay and as a zero-length exon. *Mol. Cell. Biol.* **26**, 2237–2246 (2006).
4. Shepard, S., McCreary, M. & Fedorov, A. The peculiarities of large intron splicing in animals. *PLoS ONE* **4**, e7853 (2009).
5. Thakurela, S. *et al.* Gene regulation and priming by topoisomerase II α in embryonic stem cells. *Nature Commun.* **4**, 2478 (2013).
6. Ameur, A. *et al.* Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nature Struct. Mol. Biol.* **18**, 1435–1440 (2011).
7. Rogelj, B. *et al.* Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* **2**, 603 (2012).
8. Ke, S. & Chasin, L. A. Context-dependent splicing regulation: exon definition, co-occurring motif pairs and tissue specificity. *RNA Biol.* **8**, 384–388 (2011).
9. Robberson, B. L., Cote, G. J. & Berget, S. M. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84–94 (1990).
10. McGlincy, N. J. & Smith, C. W. Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem. Sci.* **33**, 385–393 (2008).
11. Parra, M. K., Tan, J. S., Mohandas, N. & Conboy, J. G. Intraslicing coordinates alternative first exons with alternative splicing in the protein 4.1R gene. *EMBO J.* **27**, 122–131 (2008).
12. Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comp. Biol.* **11**, 377–394 (2004).
13. Jaillon, O. *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).
14. Roy, M., Kim, N., Xing, Y. & Lee, C. The effect of intron length on exon creation ratios during the evolution of mammalian genomes. *RNA* **14**, 2261–2273 (2008).
15. Pickrell, J. K., Pai, A. A., Gilad, Y. & Pritchard, J. K. Noisy splicing drives mRNA isoform diversity in human cells. *PLoS Genet.* **6**, e1001236 (2010).
16. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nature Neurosci.* **15**, 1488–1497 (2012).
17. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nature Neurosci.* **14**, 459–468 (2011).
18. King, I. F. *et al.* Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**, 58–62 (2013).
19. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. El-Andaloussi for technical support, J. Witten, J. König and Ule laboratory members for comments on the manuscript, and remaining members of the UK Brain Expression Consortium: S. Guelfi, K. D'Sa, M. Matarin, J. Vandrovcova, A. Ramasamy, J. A. Botia, C. Smith and P. Forabosco. This work was supported by the European Research Council (206726-CLIP and 617837-Translate) to J.U.; Marie Curie Post-doctoral Research Fellowship (627783-NeuroCRYSP) to L.B.; the Slovenian Research Agency (J7-5460) to J.U. and T.C.; the UK NIHR Biomedical Research Centre at Moorfields Eye Hospital and UCL Institute of Ophthalmology to V.P. and W.E.; the Wellcome Trust to S.W.W. and A.F.; the UK Medical Research Council (MRC) (U105185858) to J.U.; MRC training fellowships to C.R.S. and M.B.; and MRC project grant (G0901254), MRC training fellowship (G0802462) and MRC Sudden Death Brain Bank.

Author Contributions C.R.S., M.B. and J.U. conceived and designed the project; C.R.S., L.B., A.F., M.B., M.M. and D.T. performed experiments; C.R.S., W.E., L.B., V.P., T.C. and J.U. analysed the data and interpreted results with contributions from M.R., M.E.W. and J.H.; C.R.S. and J.U. wrote the manuscript with contributions from W.E., V.P., L.B. and S.W.W.

Author Information The sequence data and scripts are publically available from the European Genome-phenome Archive under the accession number EGAS00001001170, ArrayExpress (E-MTAB-3534), <http://icount.bioclab.si/> and https://github.com/vplagnol/recursive_splicing. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.U. (j.ule@ucl.ac.uk) and V.P. (v.plagnol@ucl.ac.uk).

METHODS

RNA-seq library preparation and sequencing. Brain samples for analysis were provided by the Medical Research Council Sudden Death Brain and Tissue Bank (Edinburgh, UK). Transcriptomic analysis of postmortem human tissue was approved by The National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint Research Ethics Committee, UK (REC reference number 10/H0716/3). All four individuals sampled were of European descent, neurologically normal during life and confirmed to be neuropathologically normal by a consultant neuropathologist using histology performed on sections prepared from paraffin-embedded tissue blocks. Twelve central nervous system regions were sampled from each individual. The regions studied were: cerebellar cortex, frontal cortex, temporal cortex, occipital cortex, hippocampus, the inferior olivary nucleus (sub-dissected from the medulla), putamen, substantia nigra, thalamus, hypothalamus, intralobular white matter and cervical spinal cord.

RNA was extracted using Qiagen tissue kits, and quality controlled as detailed previously²⁰. Libraries were prepared by the UK Brain Expression Consortium in conjunction with AROS Applied Biotechnology A/S. In brief, 100 ng total RNA was used as input for cDNA generation using the Ovation RNA-seq System V2 (NuGen Technologies). The RNA was processed according to the manufacturer's protocol resulting in amplified cDNA from total RNA and concomitant de-selection of rRNA. Notably, reverse transcription in this protocol is carried out using both oligo(dT) and random primers. This allowed total RNA profile patterns to be assessed with the latter and locations of splicing to be inferred. One microgram of the cDNA was fragmented using a Covaris S220 Ultrasonicator and the fragmented cDNA was used as the starting point for Illumina's TruSeq DNA library preparation. Finally, library molecules containing adaptor molecules on both ends were amplified through 10 cycles of PCR. The libraries were sequenced using Illumina's TruSeq V3 chemistry/HiSeq2000 and 100 base pair paired-end reads. The sequencing data was converted to fastq-files using Illumina's CASAVA Software.

RNA-seq processing. Paired end RNA-seq data was mapped to the human genome (hg19) using STAR aligner (v.2.3) with default settings and known splice junctions from GENCODE^{21,22}. For high-confidence RS-site junction detection, alignments were processed from all intronic regions >150 kb using an in-house processing pipeline implementing python (v.2.7.2), Bedtools (v.2.17.0) and R (v.3.0.0). This size limit was chosen since linear regression patterns could most readily be evaluated in such long introns (Extended Data Fig. 2a, b), and represented 943 introns in 780 genes (RefSeq release 60). Alignments from all 48 samples in >150 kb introns were combined and processed together unless indicated in the text. All spliced alignments with minimum flanking overhang of >10 nucleotides (hereafter termed anchor) and junction region exceeding 5 kb were selected and considered for further analysis. Each anchor sequence was then annotated to verify it conformed to a known splicing boundary (hereafter termed exon anchor). All further analysis was done using only those novel junctions that had a single exon anchor (Extended Data Fig. 2c). Novel junctions were then ruled out if they were not detected across either multiple brain regions or in multiple patients. We subsequently asked whether intronic sequences immediately adjacent to the novel junctions contained pentamers found at 1% of all 5' splice sites genome-wide (Extended Data Fig. 2d), or sequences located at 3' splice sites (polypyrimidine tract consisting of >11 pyrimidines present in the region of -22 to -1, including YAG as last three positions; Extended Data Fig. 2e). Novel junctions within 418 nucleotides of one another, the ninety-fifth percentile of exon lengths genome-wide, were considered in close enough proximity to have potential for exon formation. This analysis identified 2,981 novel junctions in introns >150 kb; 979 joined an upstream exon to an intronic 3' consensus splice site, 1,296 joined an intronic 5' consensus splice site with a downstream exon, and 353 pairs of junctions were proximally spaced in a manner that could form a novel exon (Supplementary Table 1, worksheet 1). For low confidence RS-site junction detection in introns >1 kb, the same process was repeated in which alignments were now processed from all intronic regions >1 kb, and the minimum novel junction span was now 100 bp. RS-sites identified in this analysis were not tested with linear regression analysis owing to shorter intron lengths having less reliable intronic read density profiles. In total, 65,173 unannotated novel junctions were detected, 43,229 of which joined intronic elements with consensus motifs of either 3' or 5' splice sites (Supplementary Table 1, worksheet 2). Of these, 40,163 were unique loci and 419 of them contained RS-site motifs. From these 419 unique and putative RS-sites, 48 were present in long gene introns (Supplementary Table 1, worksheet 1).

iCLIP library preparation, sequencing and processing. *FUS* iCLIP experiments were performed as previously described²³ with minor modifications. *FUS* iCLIP was performed with NB100-565 antibody (Novus Biologicals) at a concentration of 5 µg mg⁻¹ on human brains, while *FUS* iCLIP from mouse brain was obtained from the previous study⁷. Sequencing was performed on either an Illumina GA-II

or Illumina Miseq. The iCLIP libraries contained an experimental barcode plus a random barcode, which allowed multiplexing and the removal of PCR duplicates, respectively. The iCLIP data were mapped to hg19 using Bowtie²⁴ and further processed as described previously²³.

Computational analyses. All scripts used for the analyses in this paper are available at the Github repository (https://github.com/vplagnol/recursive_splicing).

Linear regression analysis. To establish the analysis of linear regression, each annotated intron greater than 50 kb (in at least one Ensembl transcript) was first analysed independently (Extended Data Fig. 2a, b). Following evaluation of different sized windows, we ultimately divided introns into 5-kb bins. For both the RNA-seq and *FUS* iCLIP data, we then computed the number of read pairs mapping to each bin using samtools v0.19. We then ran a regression analysis with the number of mapped reads in each bin as a dependent variable. As a test, we first used this to examine genes containing multiple introns >50 kb. This showed that slopes of fitted regression lines were comparable for different long introns of the same gene (Extended Data Fig. 2a, b). Since the slope depends on transcriptional elongation rate, this observation agrees with the finding that transcription rate is relatively constant across individual genes²⁵. We therefore assumed a constant (unconstrained) slope across each entire gene. Reducing the 5-kb bin size or the intron length cut-off reduced the reliability in the method, implying individual units of >50 kb are most appropriate for this computational analysis. Accordingly, when splitting introns into two separate parts based on novel junctions, we focused on >150 kb introns to account adequately for this size limit.

Next, for our baseline model, we coded the positions of all potential exons located in the >150 kb-intron-long gene introns (based on Ensembl annotations) using binary dummy variables and let the fitted read count data reset to an arbitrary value at each putative exon. We then considered for each intron a set of augmented models that include the same covariates at the baseline model (constant slope, dummy variable for potential exons) in addition to an additional dummy variable for each of the novel junctions identified by the split read analysis. We used a standard *F*-test *P*-value to compare the fit between the baseline model and the augmented one to quantify the improvement of the goodness-of-fit provided by each additional potential RS-site. Introns were eventually ranked on the basis of these *F*-test *P*-values, with significance threshold for further analysis set at *P* < 0.01 for both data sets (Supplementary Table 1, worksheet 3). Taken together, the following filtering workflow was used in linear regression analysis for production of Fig. 1d. (1) Select novel junctions, which connect upstream exon to deep intronic loci. Initial junctions: 1,378. (2) Exclude junctions in which the gradient remains negative after strand correction. Remaining: 1,146. (3) Select lowest *P*-value for a junction if multiple introns overlap. Remove higher *P*-values since RNA-seq has depth to identify most frequently used introns. Remaining: 536. (4) Plot after/before ratios. After/before ratios >1 correspond to increased slope, and <1 to reduced slope of linear regression line across intron. (5) Significance threshold set at *P* < 0.01 for both *FUS* and RNA-seq. Remaining: 24 junctions. (6) Select junctions with after/before ratio of >1 in both data sets. Remaining: 21 junctions representing 19 unique intronic loci; indicated by YES in column AF of Supplementary Table 1, worksheet 3.

Alternative GURAG exon analysis. All alternative exons within the UCSC Alt events track were evaluated for GURAG pentamers at their start. Two lines of evidence were then pursued to evaluate their use as RS-exons. First, we asked whether exons overlapped intronic read transition points despite being skipped. Linear regression analysis was performed on all alternative exons from UCSC Alt Events table that fell within an Ensembl transcript and would have flanking introns both >50 kb (Supplementary Table 1, worksheet 4). Analysis was performed using both RNA-seq and *FUS* data sets. Identified GURAG exons were matched to these results to determine candidate exons that show high levels of inclusion. These were subsequently followed up through evaluation of junction counts between these exons and both upstream and downstream exons within RNA-seq data, and additionally junctions between the upstream and downstream exons in which the GURAG exon would be skipped. Limited evidence for recursive splicing was considered a double-significance in linear regression analysis, but junction counts indicating that the skipped product dominates.

Second, we asked whether these GURAG exons made regular contact with upstream exons with which they are not expected to junction (based on known gene isoforms). This could indicate that the junction is used, but the GURAG exon is not included, leading to absence of isoform annotation. To identify known or novel junctions between the 99 GURAG alternative exons and upstream exons, we evaluated all junctions in RNA-seq data that were made between the identified 99 cassette exons and any annotated upstream exon (Supplementary Table 1, worksheet 5). Each junction was then enumerated and classified as 'known' or 'novel' using the known-gene UCSC annotations. If a junction was not present in this annotation database and subsequently classed as novel, then this was considered limited evidence for recursive splicing. Examples were subsequently considered

high confidence if splicing patterns inferred from the aforementioned analysis of total RNA-seq read density patterns suggested frequent use of the novel junction. Combined, these analyses identified 16 putative annotated RS-exons, two of which (in the *CADM2* and *NTM* genes) we further experimentally validate.

Cryptic element usage analysis. Introns contain numerous cryptic splicing motifs that do not produce products indicated by current gene annotations, yet might infrequently be used to create low-level alternative isoforms (Supplementary Note). To evaluate the frequency of cryptic element usage in introns of differing lengths (Extended Data Fig. 9c), we performed a search for novel junctions in our RNA-seq data that connected first intron loci to canonical second exons. Second exons were chosen because we observed that 8 out of 9 of our high-confidence sites were located in long first introns. To perform this analysis while limiting duplication of the same exon owing to multiple transcripts, RefSeq annotations were refined to include only those transcripts defined as canonical by UCSC known-gene table. Intersection of both annotation databases identified 21,531 second exons common to both databases. Of these, 798 were subsequently removed owing to a lack of evidence of gene expression across all brain regions based on gene-derived RNA-seq FPKM values. For the remaining 20,733 exons, upstream intronic regions were searched for all junctions connecting these exons to any upstream elements (Supplementary Table 3, worksheet 2). Junctions were classified according to the nature of the upstream elements. Specifically we separated into three categories; 'exon-exon' represented junctions between the canonical first exon and second exon, 'isoform' represented junctions between an alternative first exon and the second exon that are present in UCSC/RefSeq/GENCODE databases, and 'novel' represented entirely unexpected junctions between intronic elements in the UCSC/RefSeq/GENCODE databases that junction to the second exon. We restricted our final analysis of cryptic upstream elements to the 6,619 genes in which a canonical exon-exon junction was detected which accordingly span the full-length of the canonical first intron. The number of novel junctions to cryptic upstream elements was then counted in these genes, with genes grouped in bins based on the length of the canonical first intron. To avoid overlap with non-canonical minor transcripts, 'isoform' junctions were not considered. Significance between bins was determined using the Mann-Whitney *U* test with two tails.

To evaluate cryptic element usage to all 142 candidate RS-sites (high-confidence targets, all cassette exons starting with GURAG, and novel junctions detected that were consistent with RS-sites but failed to meet significance in linear regression analysis), the upstream gene body of candidate RS-sites genes were searched for all junctions present within brain RNA-seq libraries that connected these candidate RS-sites to any upstream elements (Supplementary Table 3, worksheet 3). Junctions were then classified according to the nature of the upstream elements. Specifically we asked whether the junction was to an annotated upstream exon or cryptic exon/promoter.

Gene expression comparisons. For tissue-specific gene expression comparisons in Extended Data Fig. 1, RNA-seq data from 16 human tissues obtained by the Illumina Human Body Map Project (GEO series accession number GSE30611) and RNA-seq data from 12 human tissues collected as part of the Genotype Tissue Expression (GTEx) Project (<http://www.gtexportal.org>) were mapped to hg19 genome with TopHat2²⁶. For the cell line comparisons mapped in the same way to either hg19 or mm9, data were collected from the following sources: myoblast differentiation (mm9, GEO series accession number GSE20846), erythropoiesis (hg19, GEO series accession number GSE40243) and motor neuron differentiation (mm9, GEO series accession number GSM1346027). Mean expression values across replicates was calculated using DESeq¹⁹. Tissue-specific comparisons were made between the brain and all other individual tissues for all protein coding genes. For cell-specific comparisons, differentiated cells were compared to undifferentiated cells in respective data sets. The log₂-fold expression changes were plotted as a function of gene length. In incidences in which several gene lengths were reported for a given gene, the maximum gene length was used.

Cross species intron lengths. To determine cross-species intron lengths, all human RefSeq genes were mapped to indicated species using the xenoRefGene track. Corresponding intron lengths were determined using exon start and exons end coordinates from all single-mapping transcripts. Identical introns found across multiple transcripts of the same gene were collapsed into a single unique intron for analysis so not to be counted multiple times.

GO term analysis. The GO term associated with >150 kb human UCSC genes analysed by GOrilla²⁷ using two unranked lists of genes. UCSC genes >150 kb were used as targets, while all UCSC genes were used as background. For visualization, GO terms with $>1 \times 10^{-3}$ FDR *q*-value or less than twofold enrichment were omitted.

Motif analysis. Sequence analysis around novel junction intronic loci was performed using WebLogo²⁸. Recursive exon maps were generated by string matching consensus 5' splice sites and stop codons to regions following RS-sites after con-

sidering open reading frame of upstream RefSeq exons. Strong consensus splice sites were considered GTAAG, GTGAG, GTAGG and GTATG (Fig. 2b and Extended Data Fig. 5i). Weak consensus splice sites are GTAAA, GTAAT, GTGGG, GTAAC, GTCAG and GTACG (Extended Data Fig. 5i).

Splice site score calculation. MaxEntScan was used as previously described using the First-order Markov Model setting by adding the last three nucleotides of the exon and the first 6 nucleotides of the 5' splice site¹². Competing splice site scores are presented in Supplementary Table 3, worksheet 1, and Extended Data Fig. 8d.

Conservation scores. For conservation scores, the 46-way placental mammal conservation by PhastCons track on the UCSC genome browser was used (phastCons46wayPlacental). Conservation scores were obtained for a given region using table browser, and mean scores calculated after alignment to specified features. Conservation was calculated at RS-sites ($n = 9$), at 5' splice sites downstream of RS-exons ($n = 9$), at 5' and 3' splice sites flanking constitutive exons in genes containing RS-sites ($n = 130$), and at the next two nearest 5' splice sites downstream of RS-exons ($n = 18$).

Cell culture. SH-SY5Y cells (ATTC, CRL-2266) were cultured at 37 °C, 5% CO₂ in DMEM (Life Technologies) supplemented with 10% FBS, and routinely tested for mycoplasma contamination. For all treatments in this cell line, cells were seeded to be 70–80% confluent at the day of transfection in 6-well plates.

For AON treatment, cells were transfected at 24 h with 10 μM of stated AON using Endo-porter transfection reagent (Gene-tools) as per manufacturer's instructions. At 48 h after transfection, cell media was removed and cells lysed and RNA extracted with Qiazol. All AONs were purchased from Gene-tools, and carried morpholino modifications. Sequences used to target the listed genes were: *CADM1*: AON-A1: 5'-AGCACACATGAGAAGTATGACTTAC-3'; *CADM1*: AON-A2: 5'-ATCCAAGCATAAGATTGTCACTTAC-3'; *ANK3*: AON-A1: 5'-TTTAAATGGAAAACCGCACTTAC-3'; *ANK3*: AON-A2: 5'-AATGGCC AATGCCAAGTTCCTTAC-3'.

A non-specific AON-NS that is not complementary to any locus in human genome, but has similar GC content as the AON-A1 and AON-A2, was used as control: AON-NS: 5'-CCTCTTACCTCAGTTACAATTTATA-3'.

For cycloheximide treatment after AON-A2 transfection, cells were seeded to be 50–70% confluent at the day of transfection and were treated at 48 h (first experiment) or 36 h (second experiment) with either 100 μg ml⁻¹ of cycloheximide dissolved in DMSO, or an equivalent volume of DMSO alone. At 6 h after treatment, cell media was removed and cells lysed and RNA extracted using Qiazol (Qiagen).

Zebrafish AON treatments. Zebrafish experiments were performed by injecting 1 ng of AON (Gene-tools) into the yolk of 1-cell-stage embryos. Embryos were grown at 28.5 °C and collected at 2 days post-fertilization for RNA extraction. AON NS: 5'-CCTCTTACCTCAGTTACAATTTATA-3'; AON-A1: 5'-GTGGA AAAAAATGCCAAGACTCAC-3'; AON-A2: 5'-AATGCTTCATTCACTCT GTACTCAC-3'.

Splicing reporter design. The *CADM2* splicing reporter (P1) was designed such that the RS-exon following the second *CADM2* RS-site was flanked by two short introns and the surrounding *CADM2* constitutive exons (Supplementary Table 4). Introns consisted of the first and last ~100 nucleotides of respective introns separated by multiple cloning sites. Constitutive exons were flanked by HindIII and EcoRI sites, respectively. Constructs were sub-cloned into the pcDNA3 multiple cloning site of the pBluescript plasmid using HindIII and EcoRI sites. Construct P2 was subsequently generated by removing the dominant first *CADM2* exon and first ~100 nucleotides of intron present in construct P1 with HindIII and FseI, and subcloning a separate synthetic gene product into the digested plasmid. This synthetic gene product consisted of the alternative first exon and first ~100 nucleotides of the corresponding intron. Sequences of synthetic gene products can be found in Supplementary Table 4. Mutations to both reporter variants were made by crossover PCR using construct P1 or P2 as targets and primers listed in Supplementary Table 4.

Cell fractionation. For nuclear-cytoplasmic fractionation of cell lines, samples were suspended in 1 ml cytoplasmic lysis buffer (50 mM Tris-HCl, pH 7.4, 10 mM NaCl, 0.5% NP-40, 0.25% Triton X-100, 1 mM EDTA, 1/200 volume of RNasin and 1/100 volume of protease inhibitor cocktail) and homogenized by pipetting. Sample was spun for 3 min at 3,000g. Supernatant was collected as the cytoplasmic fraction and subjected to a further spin at 10,000g for 10 min. Supernatant was removed and RNA extracted using Trizol LS (Life technologies) and the Zymogen RNAdirect extraction kit as per manufacturer's instructions. The pellet from the initial spin was retained as the nuclear fraction and lysed using Qiazol before RNA was extracted using the Zymogen RNAdirect extraction kit as per manufacturer's instructions.

RNA extraction. For cell culture experiments Qiazol (Qiagen) suspended RNA was extracted using the Zymogen RNAdirect extraction kit as per manufacturer's instructions. For brain total RNA extraction and zebrafish tissue total RNA extrac-

tion, tissue was first suspended in Qiazol (Qiagen) and homogenized using a TissueRuptor (Qiagen). RNA was then extracted using the ZymoGen RNAdirect extraction kit as per manufacturer's instructions.

RT-PCR analysis. All RNA was reverse transcribed using the high capacity cDNA synthesis kit (Applied Biosystems) using random primers and standard protocol. A total of 1 µg was used in each reaction and cDNA then diluted according to downstream application. For RT-PCR, samples were diluted 1:5 and 1 µl used for each subsequent PCR reaction. For qPCR samples were diluted 1:10 and 5 µl used for each subsequent PCR reaction.

For RT-PCR analysis, 10 ng cDNA was amplified using 2X Phusion PCR mastermix (Thermo-scientific) as per manufacturer's instructions and each primer at a final concentration of 0.5 µM. Products were run on pre-cast 6% TBE gels (Life Technologies) using low molecular mass marker (New England Biolabs) or Hyperladder V (Bioline) as a ladder. Where exon inclusion was determined from RT-PCR images, band intensity of expected product sizes were determined using ImageJ software and expressed as a percentage of total intensity for all expected bands with indicated primers.

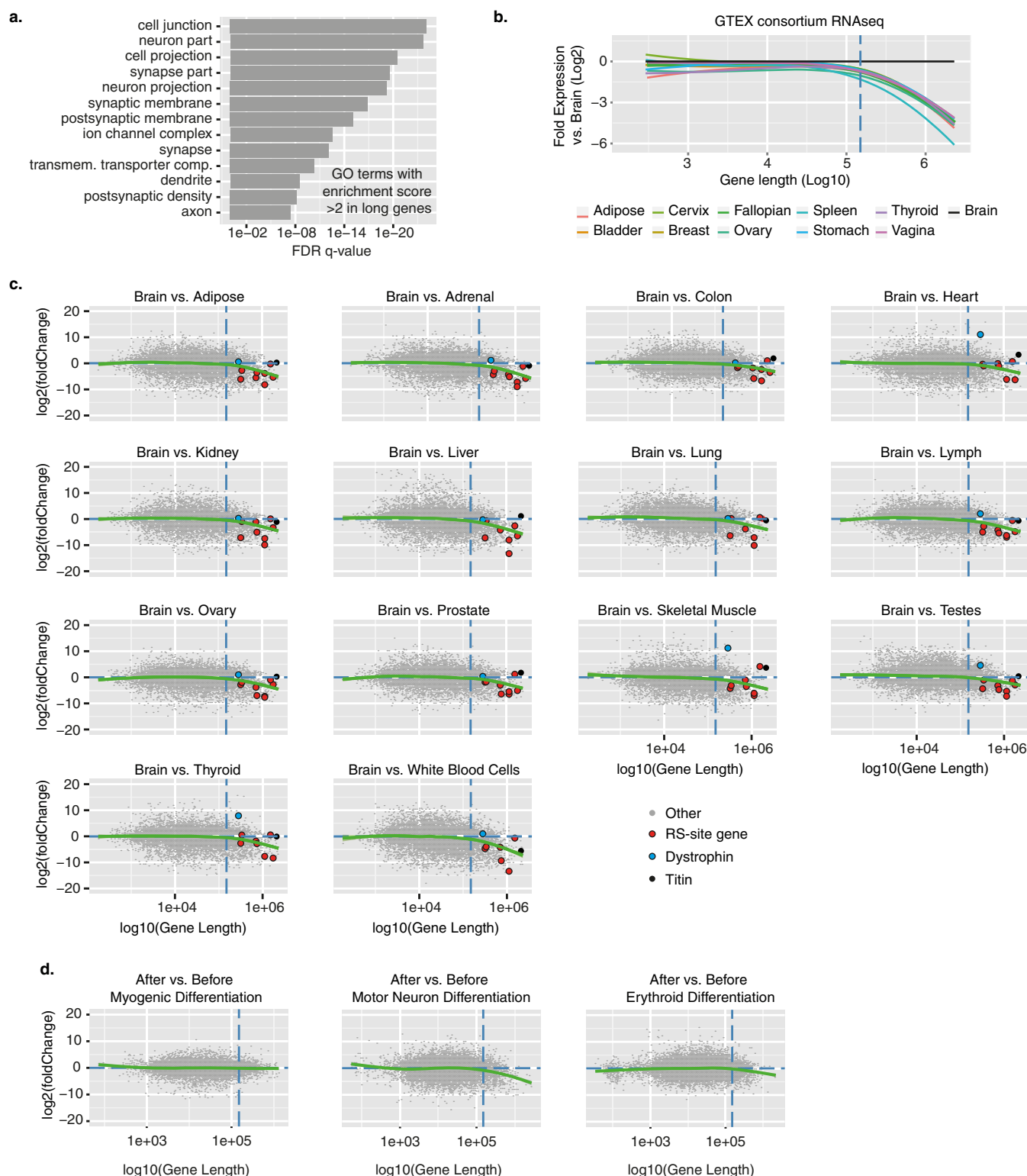
For Qiaxcel analysis cDNA was amplified with 2X Phusion PCR mastermix (Thermo-scientific) as per manufacturer's instructions and each primer at a final concentration of 0.5 µM. Samples were subsequently purified using QIAquick PCR Purification Kit and loaded onto a Qiaxcel DNA cartridge (Qiagen) and run next to a 50–800-bp DNA marker (Qiagen) on the Qiaxcel machine (Qiagen) as per manufacturer's instructions.

For qPCR analysis, 25 ng of cDNA was amplified using SYBR green PCR mastermix (Applied Biosystems) and each primer at a final concentration of 0.165 µM. PCR was carried out using an Applied Biosystems 7900HT machine as per manufacturer's instructions and quantification assessed according to standard curves generated for each primer. Signal for each interrogated junction in qPCR analysis of human genes is normalized to *GAPDH* and/or *EIF4A2* gene expression, and in zebrafish to *actb1* and *eif4a* gene expression.

Primer sequences used for RT-PCR analysis and expected product sizes can be found in Supplementary Table 4.

No statistical methods were used to predetermine sample size, and experiments were not randomized.

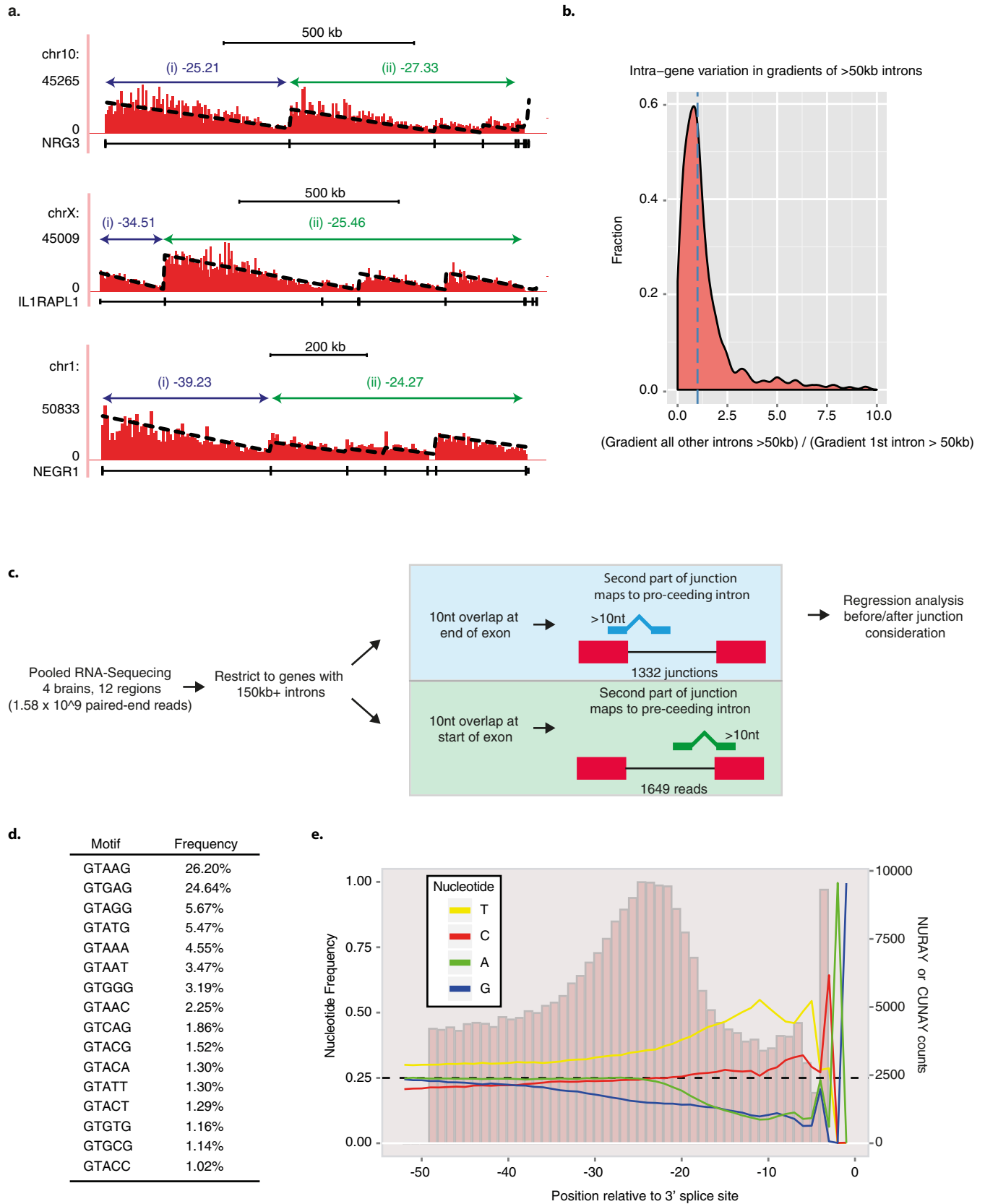
20. Trabzuni, D. *et al.* Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J. Neurochem.* **119**, 275–28 (2011).
21. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
22. Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).
23. König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Struct. Mol. Biol.* **17**, 909–915 (2010).
24. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
25. Singh, J. & Padgett, R. A. Rates of *in situ* transcription and splicing in large human genes. *Nature Struct. Mol. Biol.* **16**, 1128–1133 (2009).
26. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
27. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
28. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
29. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnol.* **28**, 511–515 (2010).
30. Herrera, F. J., Yamaguchi, T., Roelink, H. & Tjian, R. Core promoter factor TAF9B regulates neuronal gene expression. *eLife* **3**, e02559 (2014).
31. Madzo, J. *et al.* Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Rep.* **6**, 231–244 (2014).



Extended Data Figure 1 | Long gene expression is enriched in the brain.

a. GO term analysis of genes >150 kb relative to all human genes. All GO terms are associated with enrichment scores >2. **b.** The log₂-fold gene expression ratios following DESeq¹⁹ analysis of all human protein-coding genes between the brain and all other tissues. Data are represented as Loess smoothing curves after the genes by their maximum length in kilobases. Hashed vertical line indicates 150 kb gene length. RNA-seq data was obtained from the GTEx consortium. **c.** Individual scatterplots used to create Fig. 1b and representing DESeq¹⁹ analysis of individual genes within indicated tissues compared to the brain. Red dots indicate genes that contain RS-sites, blue dots indicate

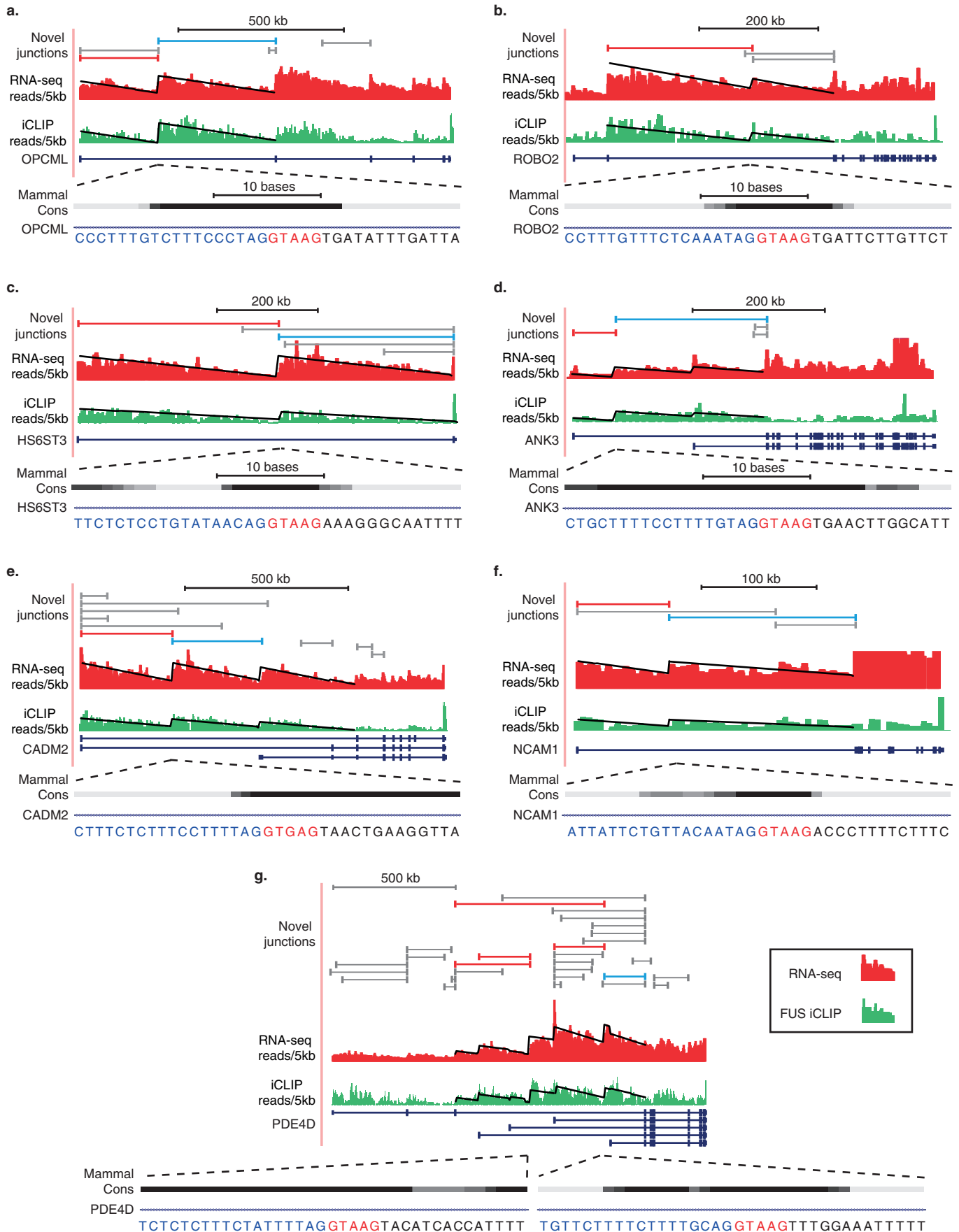
dystrophin, and black dots indicate titin (two long genes most highly expressed in muscle tissues). Grey dots are all remaining genes. **d.** DESeq¹⁹ analysis of individual gene expression after vs before differentiation of C2C12 mouse myoblasts (GSM521256) into myogenic lineage (GSM521259)²⁹, after vs before differentiation of mouse embryonic stem cells (GSM1346027) into motor neurons (GSM1346035)³⁰, or after vs before differentiation of haematopoietic stem cells (GSM992931) into erythroid lineage (GSM992934)³¹. Loess smoothing curves are shown after sorting the genes by their maximum length in kilobases. Hashed vertical line indicates 150 kb gene length.



Extended Data Figure 2 | Linear regression analysis and novel junction sequence considerations used to identify mammalian recursive splice sites.

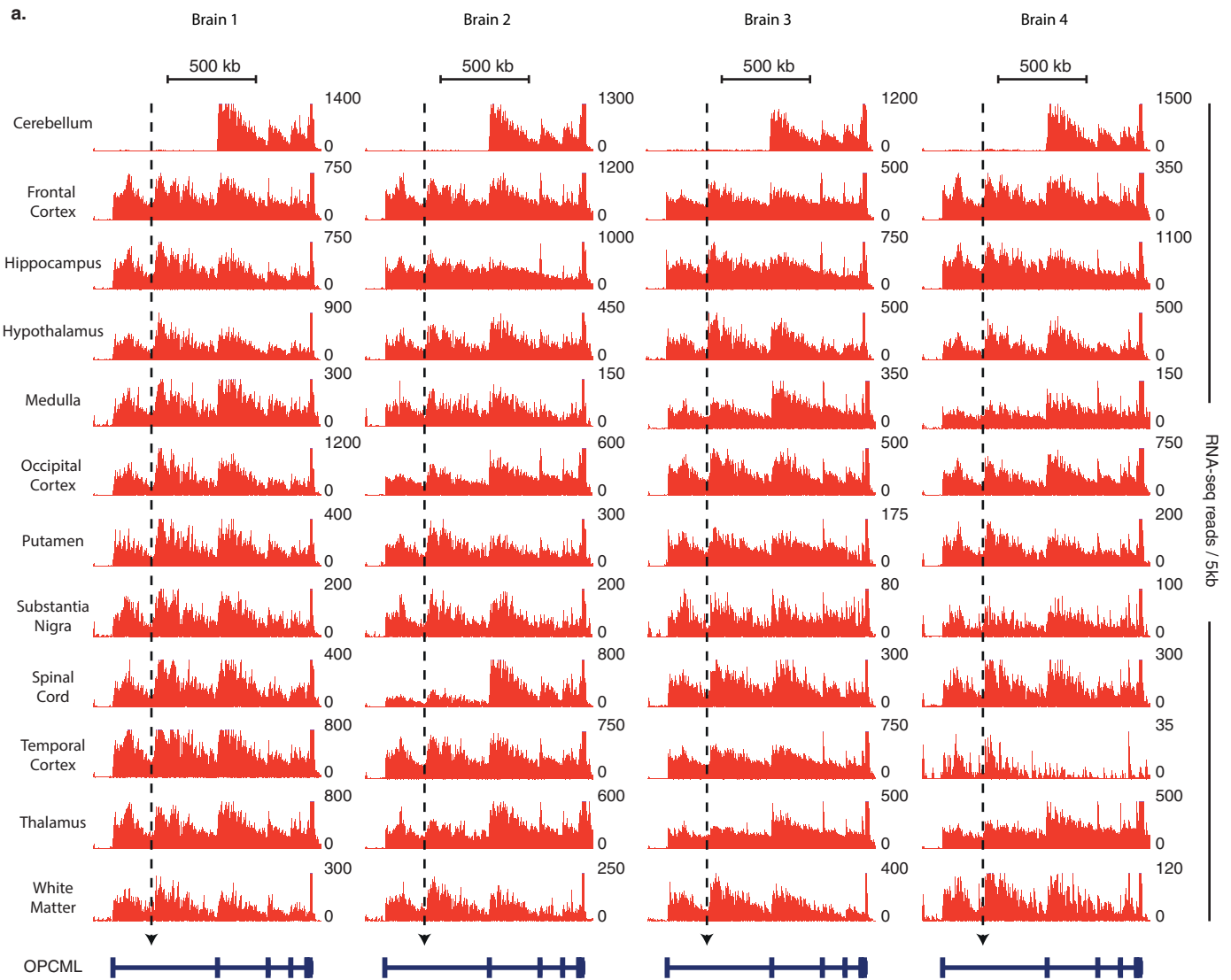
a, Examples of RNA-seq read density patterns for three genes together with their calculated gradients across the (1) first intron >50 kb, and (2) the average across all other >50-kb long introns within the same gene. Gradients represent the change in summated read count every 5 kb since RNA-seq reads are grouped in 5-kb windows and linear regression performed on resulting histograms. **b**, Density plot indicating the ratio of gradients of all other >50 kb

introns within the same gene: the gradient of the first intron >50 kb. Blue hashed line represents ratio of 1. This would indicate that gradients for long introns within the same gene are comparable and transcription is proceeding at a largely constant rate. **c**, Schematic of the bioinformatics pipeline used to identify novel junctions. **d**, Ranking of human 5' splice site pentamer usage genome-wide. **e**, Nucleotide usage frequency at human 3' splice sites genome-wide, and branch-point positioning relative to 3' splice site genome-wide.



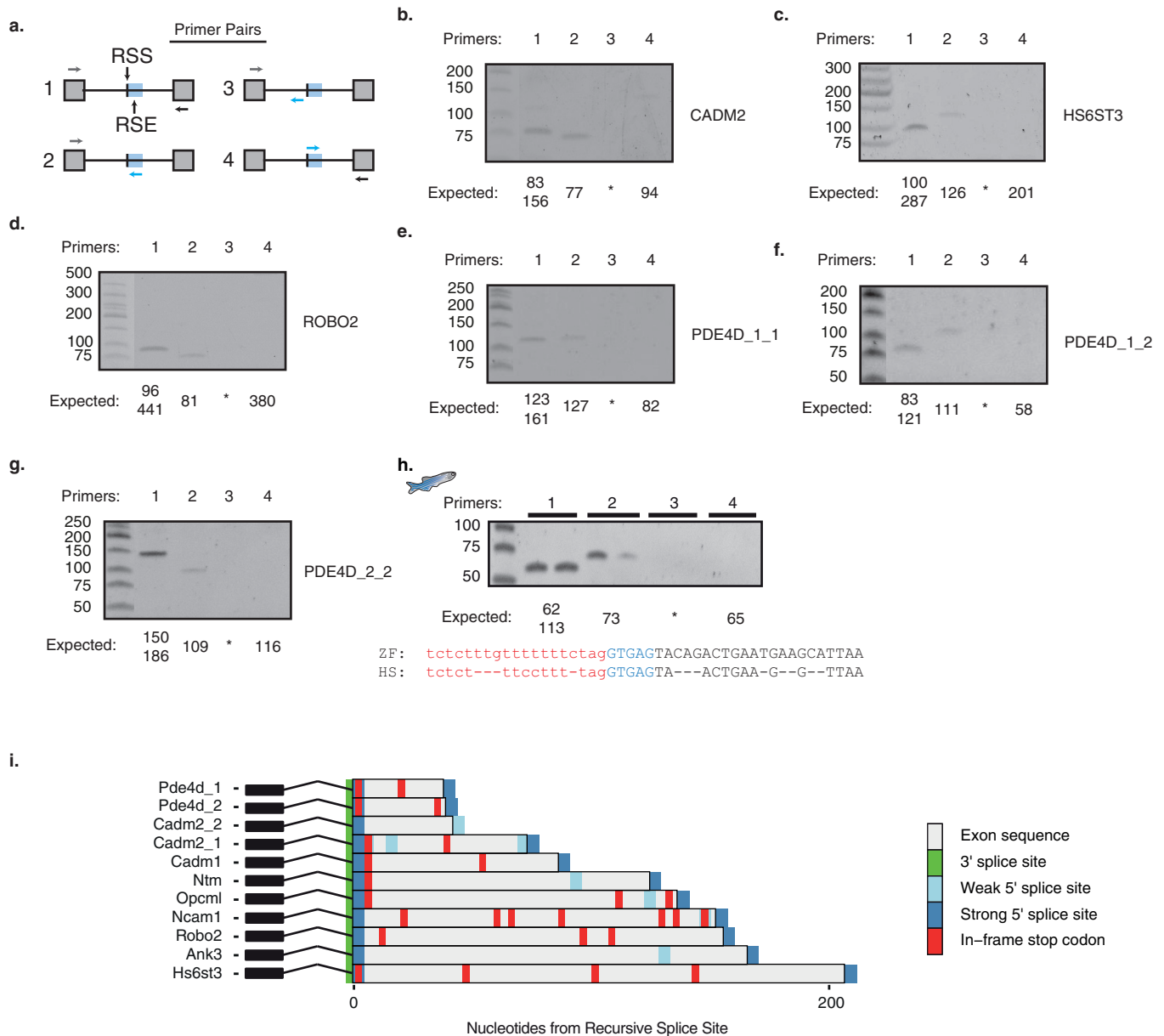
Extended Data Figure 3 | Inferred splicing patterns identify recursive splice sites within mammalian > 150 kb intron genes. **a–g**, RNA-seq (red) read density patterns and normalized *FUS* iCLIP (green) cross-link density patterns for the *OPCML* (**a**), *ROBO2* (**b**), *HS6ST3* (**c**), *ANK3* (**d**), *CADM2* (**e**), *NCAM1* (**f**) and *PDE4D* (**g**) genes within human brains. RNA-seq reads and normalized *FUS* iCLIP cross-links are grouped in 5-kb windows. RefSeq introns >150 kb were searched for novel junctions and linear regression performed on all Ensembl introns >50 kb in which novel junctions were located. Gene isoforms

displayed are those including introns within which significant junctions were identified. Red novel junctions represent significant improvements in goodness-of-fit in both RNA-seq and *FUS* regression analysis ($P < 0.01$ in both data sets, *F*-test). Blue novel junctions contact RS-exons. Grey novel junctions were not deemed significant following regression analysis. Zoomed area represents sequence at deep intronic loci surrounding novel junction. Phylo-P conservation track indicates sequence conservation across 46 levels of mammalian evolution.



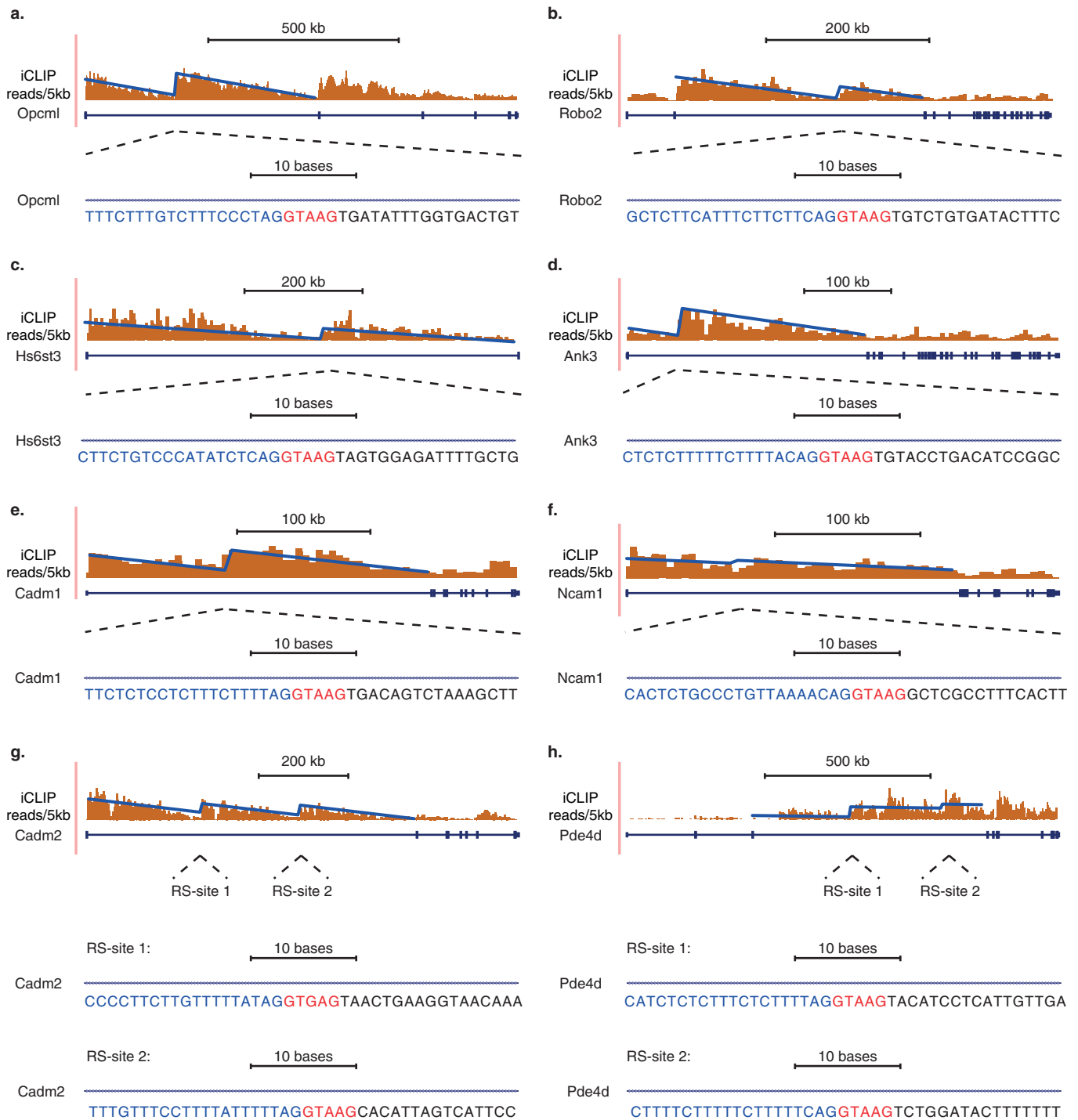
Extended Data Figure 4 | Inferred recursive splicing patterns in the *OPCML* gene across four separate brains. **a**, RNA-seq read density patterns for the *OPCML* gene across 12 different regions of four separate brains. Gene isoform

displayed is that which included the long first intron within which a significant novel junction was identified. RNA-seq reads are grouped in 5-kb windows. Dotted arrows indicate location of experimentally derived RS-site.



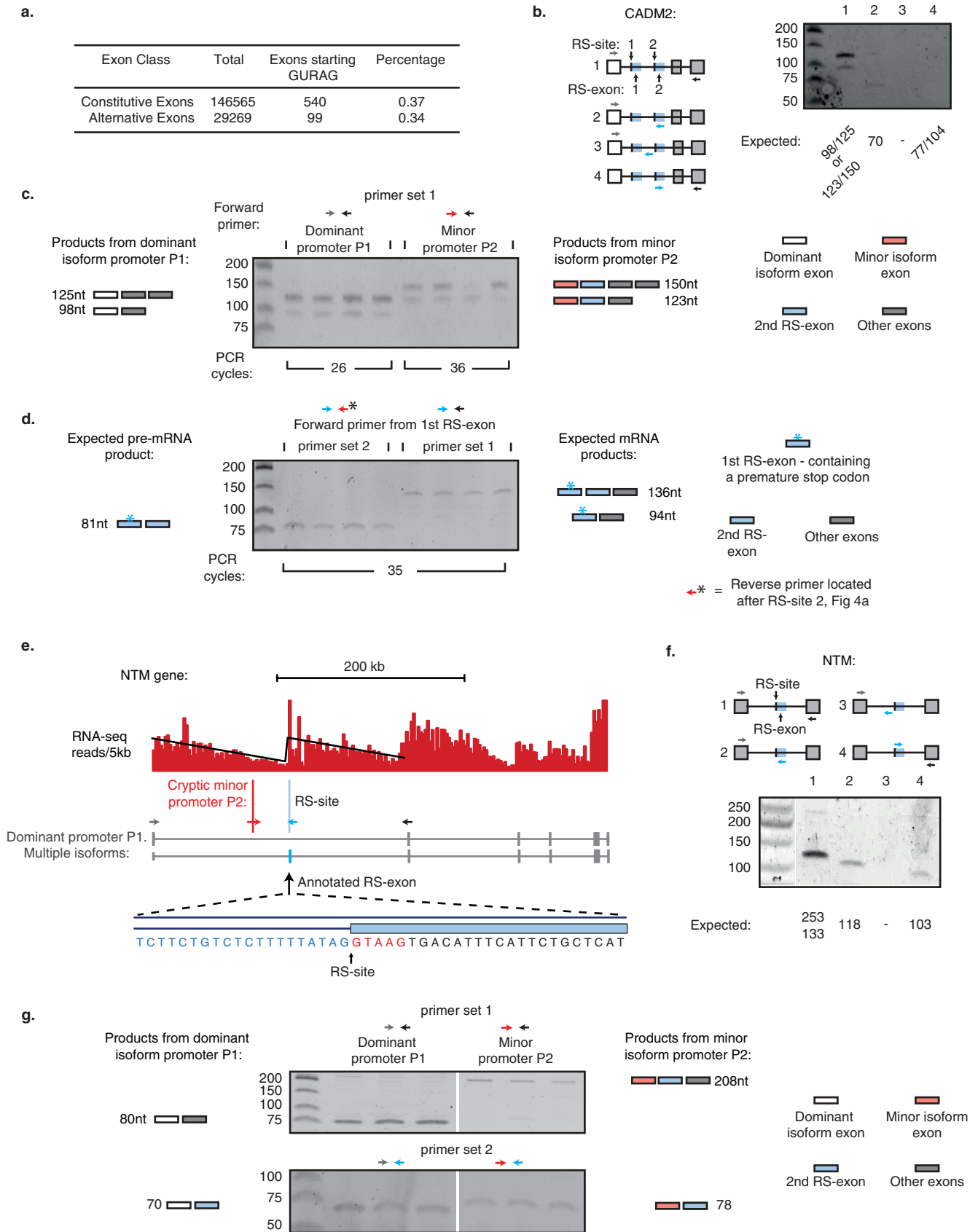
Extended Data Figure 5 | RT-PCR confirmation of RS-sites in human and zebrafish samples, and prediction of mouse RS-exons. **a**, Schematic of primer design used for RT-PCR validation of novel junctions. **b–g**, RT-PCR analysis of *CADM2* (**b**), *HS6ST3* (**c**), *ROBO2* (**d**), *PDE4D_1_1* (**e**), *PDE4D_1_2* (**f**) and *PDE4D_2_2* (**g**) genes around RS-sites using indicated primers. For *PDE4D* sites, first number after gene name indicates RS-site studied, second number indicates the upstream exon used. See Extended Data Fig. 3g for junctions

detected. **h**, RT-PCR analysis of *cadm2a* RS-site junction in adult male and female zebrafish embryos, together with an alignment of zebrafish (ZF) *cadm2a* RS-site to human (HS) *CADM2* RS-site. **i**, Map of consensus splice site location and in-frame termination codons following RS-sites in indicated mouse genes. Strong consensus splice sites are GTAAG, GTGAG, GTAGG and GTATG. Weak consensus splice sites are GTAAA, GTAAT, GTGGG, GTAAC, GTCAG and GTACG.



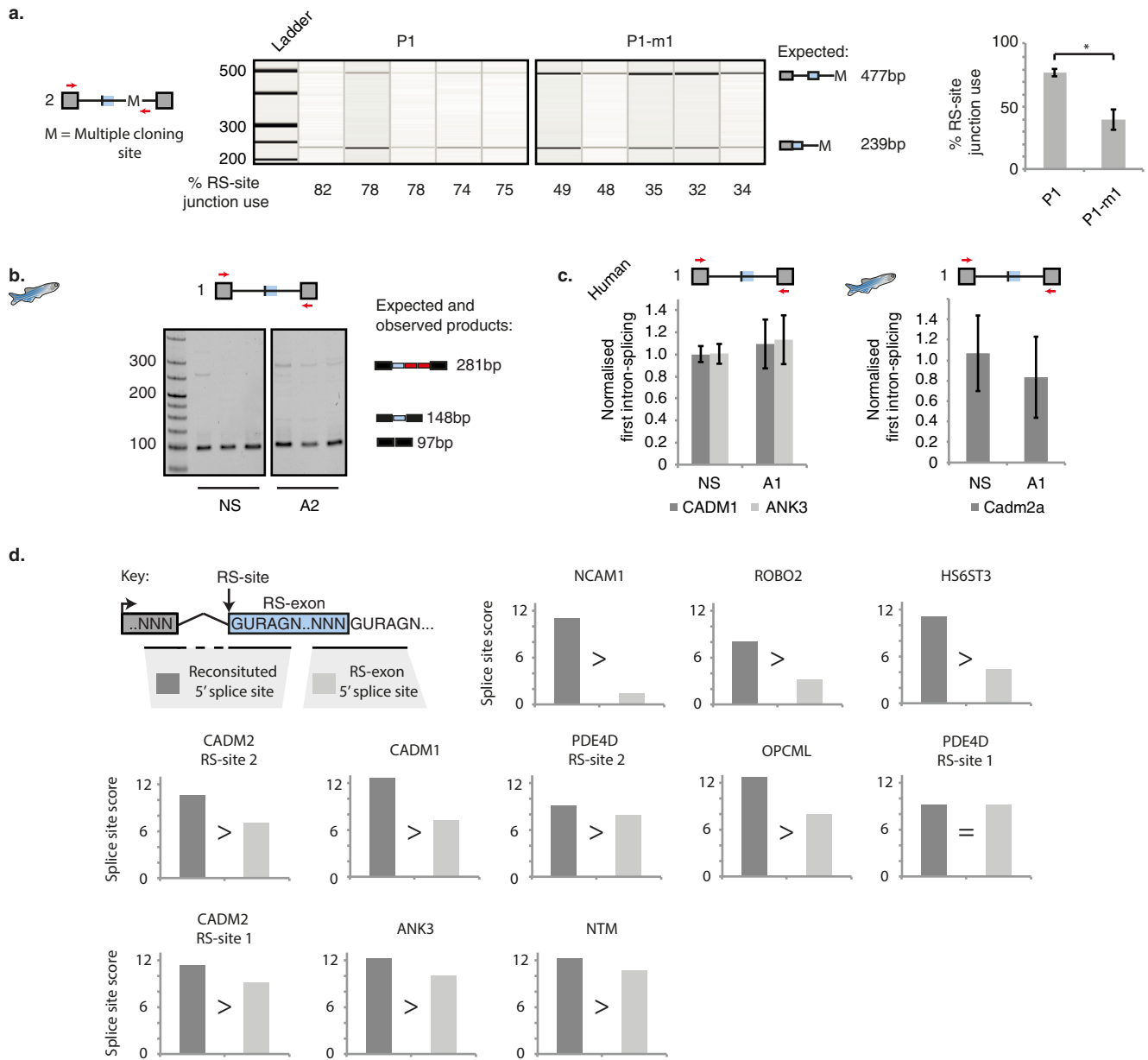
Extended Data Figure 6 | Conservation of inferred recursive splicing patterns in the mouse brain. a–h, Normalized *Fus* iCLIP read density patterns for the *Opcml* (a), *Robo2* (b), *Hs6st3* (c), *Ank3* (d), *Cadm1* (e), *Ncam1* (f), *Cadm2* (g) and *Pde4d* (h) genes within the mouse brain. Normalized *FUS*

iCLIP cross-link sites are grouped in 5-kb windows, and the displayed linear regression lines were computed on resulting histograms. Zoomed area at deep intronic loci represents RS-site sequences conserved from humans to mouse.



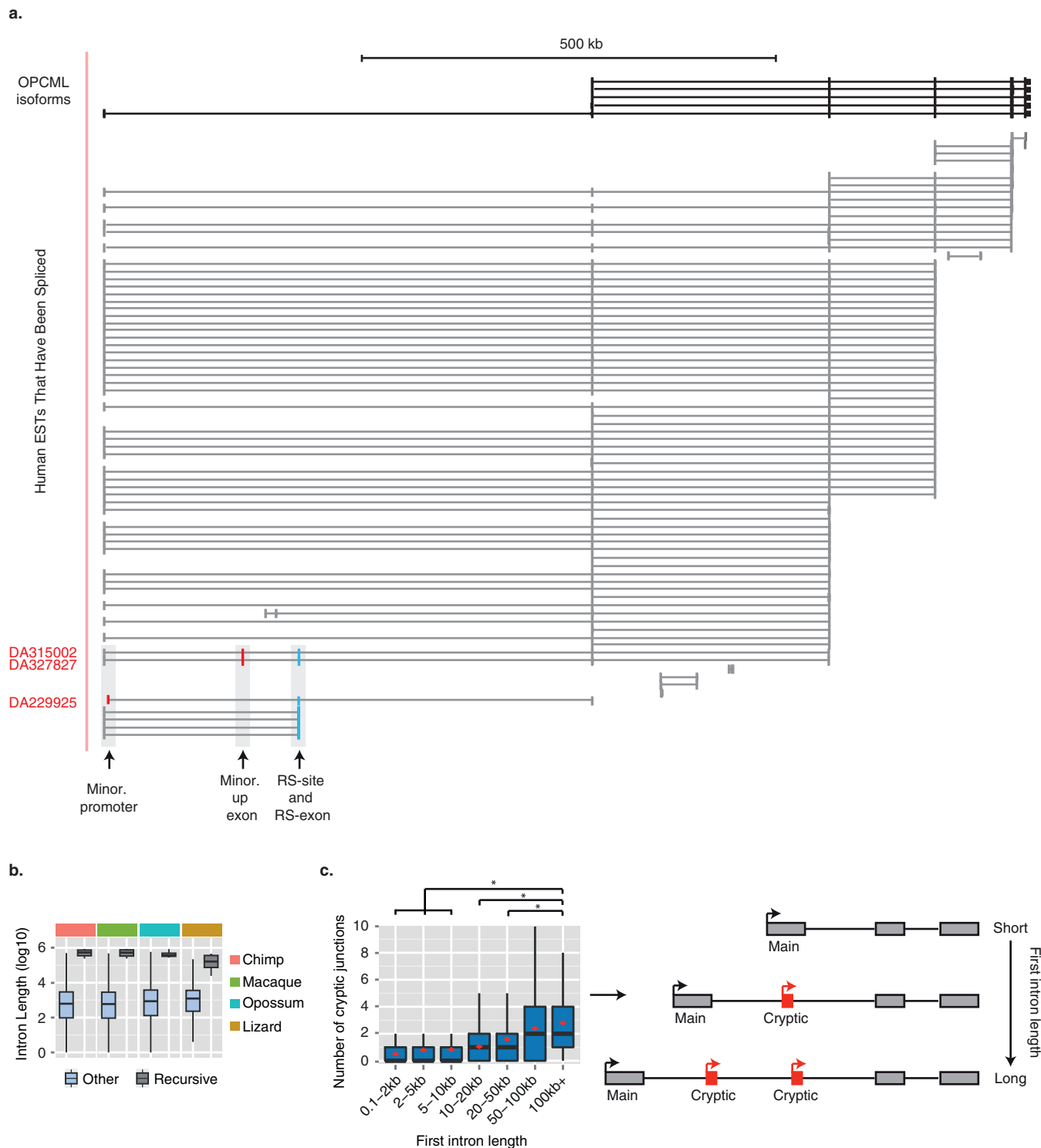
Extended Data Figure 7 | Promoter-dependent inclusion of RS-exons in *CADM2* and *NTM* genes. **a**, Number of cassette and constitutive exons starting with motif GURAG. **b–d**, RT-PCR of *CADM2* gene in the frontal cortex using primers indicated in **b** or Fig. 4a. RT-PCR was carried out on one (**b**) or four (**c**, **d**) human brains. In **c**, the inclusion of the second RS-exon occurs together with the minor promoter. Two bands are present for both PCR reactions due to the presence of an alternatively spliced exon following the RS-exon. This can result in two distinct long or short isoforms. In **d**, the inclusion of the second RS-exon occurs when the first RS-exon is included. Schematics in **c** and **d** represent examined splicing products together with expected length of products. **e**, RNA-seq read density patterns for the *NTM* gene and expected human isoforms. RNA-seq reads are grouped in 5-kb windows and linear

regression performed on resulting histograms. A cryptic minor promoter/exon detected by RNA-seq is indicated by vertical red line. The annotated RS-exon is indicated by the vertical blue line. Zoomed area represents RS-site sequence at start of the annotated RS-exon. Primers to assess the major and minor promoter products associated with the RS-exon are indicated by coloured arrows. **f**, RT-PCR of *NTM* gene around RS-exon using indicated primers. **g**, RT-PCR analysis of *NTM* products in which the upstream exon is either derived from the major upstream promoter or the cryptic upstream promoter/exon. RT-PCR was performed in the frontal cortex of three human brains using primer sets indicated by coloured arrows in **e**. Schematics represent possible splicing products together with expected length of products. Top panel assesses RS-exon inclusion, bottom panel assesses RS-site junction detection.



Extended Data Figure 8 | Recursive splicing regulates the alternative splicing of RS-exons. **a**, Qiagel analysis and quantification of the splicing intermediates of indicated *CADM2* splicing reporter products following transfection in SH-SY5Y cells. Primers used are indicated by red arrows in schematic, together with expected products and their sizes. **b**, RT-PCR analysis of the zebrafish *cadm2a* mRNA after *in vivo* injection of AON-2. Sequencing reveals RS-exon inclusion results in subsequent splicing to additional

downstream cryptic elements before the second exon, explaining why RS-exon included product size is larger than expected. **c**, qRT-PCR analysis of exon-exon junctions surrounding the RS-site containing introns following AON-A1 mediated inhibition of RS-site use of the human *CADM1* and *ANK3* genes ($n = 3$, 1 experiment) or the zebrafish *cadm2a* gene ($n = 7$, 3 separate experiments). **d**, Splice site scores of reconstituted 5' splice sites following first step of recursive splicing versus the 5' splice sites of corresponding recursive exons.



Extended Data Figure 9 | Cryptic elements are frequent in long first introns.

a. UCSC annotated isoforms of the *OPCML* gene together with spliced expressed sequence tags (ESTs) detected across the *OPCML* locus. Recursive exon is marked in blue, and the preceding exons produced by minor promoter or cryptic splicing of the long first intron are marked in red. **b.** Lengths of the 9 introns containing the high-confidence RS-sites compared to other introns across vertebrates. Results are an extension of Fig. 4g. **c.** Boxplot showing the detected number of unannotated alternative start exons that junction to the

dominant second exon of brain expressed genes. Only novel junctions that do not match UCSC/GENCODE transcripts are considered for analysis. Genes are separated into bins based on the first intron length of the canonical isoform. Boxplot presents median, first and third quartile boundaries for each bin. Additional red diamonds indicate mean values for each bin. $*P < 10^{-10}$ (Mann-Whitney *U* test). Only tests between the 100 kb+ bin to other bins are shown. Right panel shows cartoon of the implications of boxplot results.

Genome-wide identification of zero nucleotide recursive splicing in *Drosophila*

Michael O. Duff^{1*}, Sara Olson^{1*}, Xintao Wei^{1*}, Sandra C. Garrett¹, Ahmad Osman¹, Mohan Bolisetty¹, Alex Plocik¹, Susan E. Celniker² & Brenton R. Graveley¹

Recursive splicing is a process in which large introns are removed in multiple steps by re-splicing at ratchet points—5' splice sites recreated after splicing¹. Recursive splicing was first identified in the *Drosophila Ultrabithorax* (*Ubx*) gene¹ and only three additional *Drosophila* genes have since been experimentally shown to undergo recursive splicing^{2,3}. Here we identify 197 zero nucleotide exon ratchet points in 130 introns of 115 *Drosophila* genes from total RNA sequencing data generated from developmental time points, dissected tissues and cultured cells. The sequential nature of recursive splicing was confirmed by identification of lariat introns generated by splicing to and from the ratchet points. We also show that recursive splicing is a constitutive process, that depletion of U2AF inhibits recursive splicing, and that the sequence and function of ratchet points are evolutionarily conserved in *Drosophila*. Finally, we identify four recursively spliced human genes, one of which is also recursively spliced in *Drosophila*. Together, these results indicate that recursive splicing is commonly used in *Drosophila*, occurs in humans, and provides insight into the mechanisms by which some large introns are removed.

Recursive splicing was first identified in the *Drosophila melanogaster* *Ultrabithorax* (*Ubx*) gene¹. The 73-kb intron within *Ubx* contains two alternative microexons (mI and mII) that both contain the consensus 5' splice site sequence GTAAGA immediately downstream of the 3' splice sites. In addition, this intron contains a ratchet point, a zero nucleotide exon consisting of juxtaposed 3' and 5' splice sites. It has been shown¹ that rather than being removed in a single step, the 73-kb *Ubx* intron is removed in four steps in which the upstream constitutive exon is spliced to exon mI, and subsequently re-spliced to exon mII, the ratchet point, and finally the downstream constitutive exon. A previous genome-wide computational search for potential ratchet points conserved between *D. melanogaster* and *D. pseudoobscura* predicted 160 potential ratchet points in 124 introns of 106 genes². Of these, only seven ratchet points in three genes (*kuzbanian* (*kuz*), *outspread* (*osp*) and *frizzled* (*fz*)) have been reported to be experimentally validated^{2,3}.

We generated 10.9 billion uniquely mapped reads of rRNA-depleted, paired-end, strand-specific RNA sequence from 183 *D. melanogaster* individual RNA samples comprising 35 dissected tissue samples, 24 untreated and 11 ecdysone treated cell lines, 30 distinct developmental stages and males and females of four strains from the *D. melanogaster* Genetic Reference Panel⁴ (Supplementary Table 1). The majority of these RNA samples were previously used to generate poly(A)⁺ RNA sequence data^{5,6}. As the current libraries were prepared without poly(A) selection, they contain a mixture of mRNA, pre-mRNA and nascent RNA. Co-transcriptional splicing can be observed in total, nuclear, or nascent RNA-seq data by the saw-tooth pattern of repeatedly decreasing read density across introns in the 5' to 3' direction of transcription⁷ (Fig. 1a). While visually inspecting these data on a genome browser, we noticed several large introns that lacked

internal annotated exons yet possessed saw-tooth patterns of read density suggestive of co-transcriptional splicing, including the introns from *Ubx* (Fig. 1b), *kuz*, *osp* and *fz* that were previously shown to undergo recursive splicing. We hypothesized that such saw-tooth patterns could be indicative of recursive splicing, and therefore performed a genome-wide search for ratchet points supported by the RNA-seq data.

To identify potential zero nucleotide exon-type ratchet points, we parsed the RNA-seq alignments to identify novel splice junctions where the reads mapped to an annotated 5' splice site and an unannotated 3' splice site, and the genomic sequence at the 3' splice site junction was AG/GT (Extended Data Fig. 1a). We also aligned the total RNA-seq data to a database of splice junctions between annotated exons and all potential ratchet points (AG/GT sequences) in the downstream intron that did not correspond to annotated 3' splice sites. We then identified ratchet point junctions where reads mapped without any mismatches, with at least three distinct offsets, and with an overhang of at least eight nucleotides (Extended Data Fig. 1b). We then visually inspected each ratchet point independently identified by both methods on the genome browser, removing candidates that did not display an obvious saw-tooth pattern of read density or which clearly corresponded to an unannotated exon.

We identified a total of 197 ratchet points in 130 introns of 115 genes (Supplementary Table 2). Two of these ratchet points were missed by our computational approaches, but identified during the course of manual inspection on the browser, validated, and subsequently included in the remainder of these analyses. This provides the first experimental verification of 91 of the 160 ratchet points computationally predicted by ref. 2 based on comparative genomics (Supplementary Table 3). Of the 69 unverified ratchet points predicted by ref. 2, 34 correspond to previously unannotated exons, 23 lacked convincing saw-tooth patterns, 7 did not pass our recursive junction thresholds, and 5 could not be identified in the current assembly of the genome. Although it is difficult to conclude that these are not true ratchet points, we have not included them in our subsequent analysis as their supporting evidence is inconclusive. The other 106 (53.8%) of the ratchet points we identified are described here for the first time.

Most genes (100) contain only one recursively spliced intron, although 15 genes contain two. The number of ratchet points in an intron ranges from one to six (Extended Data Fig. 2a). The recursively spliced introns range in size from 11,341 bp to 132,736 bp, with an average size of 45,164 bp. Recursive-splice-site-containing introns are enriched in large introns (97% of all introns are smaller than the smallest recursive intron), although not all large introns contain recursive splice sites (Extended Data Fig. 2b). In fact, only 6% of introns larger than the smallest recursive intron are recursively spliced. The segments of the introns removed by recursive splicing range from 2,596 bp to 63,580 bp, with an average size of 17,953 bp ($\pm 9,039$ bp) and median size of 16,368 bp (Extended Data Fig. 2c). The *luna* gene contains a 108-kb intron with five ratchet points, such that the intron is

¹Department of Genetics and Genome Sciences, Institute for Systems Genomics, University of Connecticut Health Center, Farmington, Connecticut 06030, USA. ²Department of Genome Dynamics, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

*These authors contributed equally to this work.

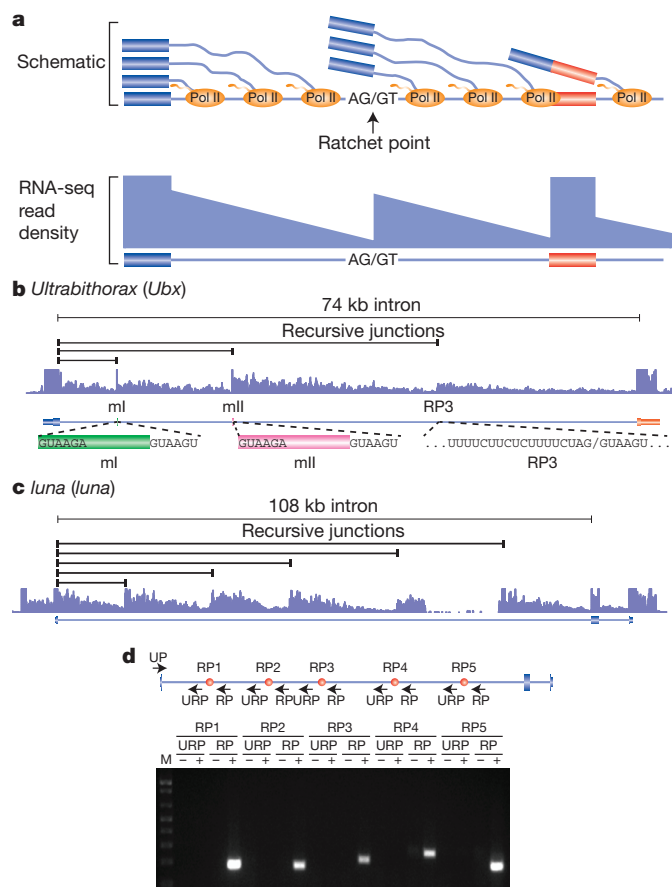


Figure 1 | Identification and validation of recursive splice sites in *Drosophila*. **a**, Schematic diagram of nascent pre-mRNA transcripts during co-transcriptional splicing and the corresponding read density that would be observed in total RNA-seq data. Note the saw-tooth pattern created by the 5' to 3' gradient of RNA-seq read density from the exon to the downstream ratchet point and splice site. **b**, Example of total RNA-seq data for the *Ubx* gene which is known to contain three recursive splice sites. Also shown are the splice junction reads supporting recursive splicing at each site. **c**, Example of five recursive splice sites identified in *luna*. Shown are the recursive junctions identified and the overall RNA-seq read density from all samples (blue). **d**, RT-PCR validation of the *luna* ratchet points (red dots) using primers in the upstream constitutive exon and flanking the putative ratchet points (UP). The RP primers are expected to yield RT-PCR products if the constitutive exon is spliced to the ratchet point. The URP primers, which are upstream of each ratchet point, serve as negative controls.

removed in six stepwise recursive splicing events (Fig. 1c). The five ratchet points are supported by the saw-tooth pattern of read density across the intron, reads that map to the exon-ratchet point splice junctions (Fig. 1c), and have been validated by RT-PCR and Sanger sequencing (Fig. 1d). In total, RT-PCR and Sanger sequencing validated 24 ratchet points from 14 genes in *Drosophila* S2 cells (Extended Data Fig. 3).

Ratchet points are zero nucleotide exons, and therefore do not exist in the mRNA. However, direct evidence of recursive splicing can be obtained by identifying lariat introns—by-products of all splicing reactions that contain a 2'-5' linkage between the first nucleotide of the intron and the branch point. Because reverse transcriptase can occasionally traverse the branch point, reads corresponding to the 5' splice site-branch-point junction may be present in the total RNA-seq data (Fig. 2a). To identify putative recursive lariat introns, we generated a set of potential 5' splice site-branch-point junctions for all recursively spliced introns, and all possible permutations, and aligned the total RNA-seq reads to them (Methods). Although rare, we identified 46

reads that mapped uniquely to 27 recursive lariat introns in 20 genes (Supplementary Table 4). Directed RT-PCR and sequencing experiments independently verified 14 recursive lariats in 9 genes (Extended Data Table 1) for a total of 41 recursive lariat introns in 26 genes. Ten of the lariat introns detected correspond to the first segment of the recursive introns and are also supported by standard splice junction reads. However, the remaining lariat introns detected correspond to internal segments, further supporting the sequential nature of recursive splicing. For example, *couch potato* (*cpo*) contains an intron that is removed in three recursive splicing events involving two ratchet points. We obtained evidence for all three lariats from both the total RNA-seq data and directed RT-PCR sequencing experiments (Fig. 2b). This analysis also identified the putative branch points used for these recursive splicing events. All but five of these branch points reside from -42 to -19 upstream of the 3' splice site with a peak at -29 (Fig. 2c). Six of the 3' splice sites appear to use two different branch points. We observed that 81% have an A at the branch point, while 12%, 5% and 2% have a T, C, or G, respectively (Fig. 2d).

The nucleotide sequences of ratchet points resemble juxtaposed 3' and 5' splice sites (Fig. 3a), and the regions immediately flanking the ratchet points are much more highly conserved than those flanking non-ratchet point AG/GT sequences in the same introns (Fig. 3b). However, the ratchet points have a more prominent pyrimidine tract, and a significantly ($P \leq 0.0001$) higher frequency of a TT dinucleotide at positions -5 and -6 relative to the 3' splice site when compared to introns genome-wide. Whereas only 43.76% (30,151 out of 68,898) of all introns have Ts at positions -5 and -6, 99.5% (196/197) of ratchet points do. The only ratchet point lacking a TT dinucleotide at positions -5 and -6 is in *CG15360* which has a C at position -6 that is conserved in other *Drosophila* species. Notably, the majority of *Caenorhabditis elegans* 3' splice sites have this sequence⁸ and it has been shown that the large U2AF subunit (encoded by *U2af50* in *Drosophila*) interacts with these bases. Thus, the strong preference for the TT dinucleotide at positions -5 and -6 of *Drosophila* ratchet points could represent high affinity U2AF binding sites so that the ratchet points are efficiently recognized.

To test this hypothesis, we sequenced total RNA from untreated S2 cells as well as cells treated with double-stranded RNA (dsRNA) to knockdown expression of *lariat debranching enzyme* (*ldbr*) as a control, *U2af38*, *U2af50*, or both *U2af38* and *U2af50* (Extended Data Table 2). We observed ~20 recursive junction reads per million mapped reads that mapped to 119 and 100 distinct ratchet points in untreated controls and *ldbr* depleted cells, respectively (Fig. 3c). Depletion of *U2af38* or *U2af50* alone reduced the frequency of recursive junction reads 3–4-fold (corresponding to 81 and 64 distinct ratchet points) (Fig. 3c). Notably, depletion of both *U2af38* and *U2af50* resulted in a complete absence of detectable recursive junction reads (Fig. 3c), although similar fractions of non-recursive junction reads were observed in all samples (Fig. 3d). Depletion of U2AF might have such a strong impact on recursive splicing—but not on non-recursive splicing—because recursive junction reads can only be generated from nascent RNA while non-recursive junction reads can be generated from stable mRNAs. Additionally, exon or intron definition may not be possible for zero nucleotide exons in *Drosophila* due to the combination of large introns and non-existent exons. This would eliminate many of the cooperative interactions normally involved in splice site recognition, making recursive splicing particularly sensitive to decreased U2AF levels. Although additional work will be required to fully elucidate the role for U2AF in recursive splicing, this result strongly suggests that U2AF is required for efficient recognition of ratchet points in S2 cells.

To determine whether recursive splicing is evolutionarily conserved, we generated rRNA-depleted, stranded RNA-seq data from *Drosophila simulans*, *Drosophila sechellia*, *Drosophila yakuba*, *Drosophila pseudoobscura* and *Drosophila virilis* adults (Extended Data Table 3). We aligned these data to the corresponding reference genomes and searched for splice junction reads whose 3' splice sites

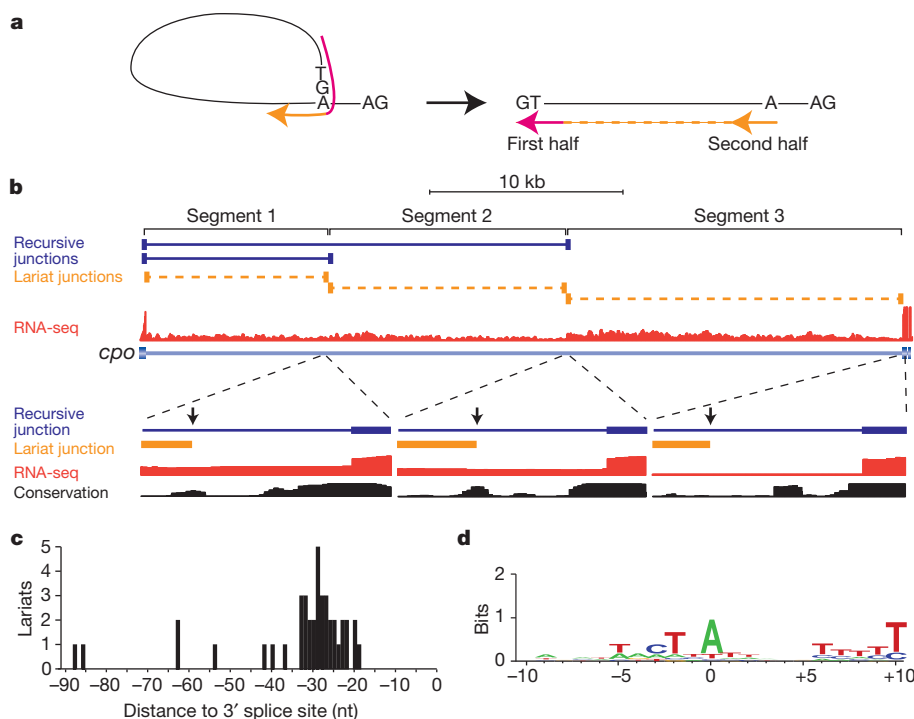


Figure 2 | Identification of recursive lariet introns in *Drosophila*. **a**, RNA-seq reads (red and orange indicate the first and second half of an individual read) that traverse a 5' splice site–branch-point junction would align to the linear intron as out-of order split reads. **b**, Example of recursive lariet introns in *cpo*. Shown are the recursive junctions identified (blue), the lariet junction reads (orange) and the overall RNA-seq read density from all samples (red). A magnification of each branch-point region is also shown along with the conservation among 16 insects. The positions of the branch points are indicated by the vertical arrows. **c**, Distribution of the distance of the recursive lariet intron branch points from the 3' splice sites. **d**, Sequence logo of the recursive lariet intron branch-point sequences.

mapped to positions orthologous to the identified *D. melanogaster* ratchet points. Despite having two orders of magnitude fewer reads from these species, 131 of the 197 *D. melanogaster* ratchet points (66.5%) were identified in at least one of the five other *Drosophila* species, 69 of which were identified in at least two species (Extended Data Table 4 and Supplementary Table 5). Together, these observations demonstrate that the nucleotide sequence and function of ratchet points are conserved among *Drosophila* species, indicating that recursive splicing is evolutionarily conserved.

We also searched for zero nucleotide exons in humans by generating and analysing >1.1 billion reads of total RNA data from 20 tissues (Extended Data Table 5). Our analysis pipeline identified 76 putative ratchet points, but upon further inspection all but five ratchet points in four genes were eliminated because they either lacked an obvious saw-tooth pattern of read density or corresponded to unannotated exons. The five ratchet points that we identified in *PDE4D*, *HS6ST3*, *CADM2* and *ROBO2* (Supplementary Fig. 1 and Supplementary Table 6) were independently identified by an accompanying study⁹ that also demonstrated that recursive splicing in humans involves recursive exons rather than true zero nucleotide exons as in *Drosophila*. This suggests that although recursive splicing occurs in both *Drosophila* and humans, the precise nature of recursive splicing differs between these organisms. Nonetheless, the presence of ratchet points in both *Drosophila* *Hs6st* and its human orthologue *HS6ST3* indicates that recursive splicing is either very ancient or evolved independently.

The host genes containing recursively spliced introns are expressed in a broad spectrum of developmental time points, tissues and cell types—the recursive host genes are expressed at fragments per kb per million mapped reads (FPKM) >1 in 72%, 93% and 83% of cell lines, developmental time points, and tissues, respectively. However, host gene expression levels are quite dynamic throughout development and 63% have their peak expression in nervous system tissues (Fig. 4a), consistent with Gene Ontology (GO) enrichments in development and neural functions (Supplementary Table 7).

Several lines of evidence suggest that recursive splicing is constitutive—specifically, when the host gene is transcribed, it is recursively spliced. First, we have been unable to detect lariet introns that would be generated by ratchet point skipping or the direct splicing of the flanking constitutive exons without recursive splicing. In our directed

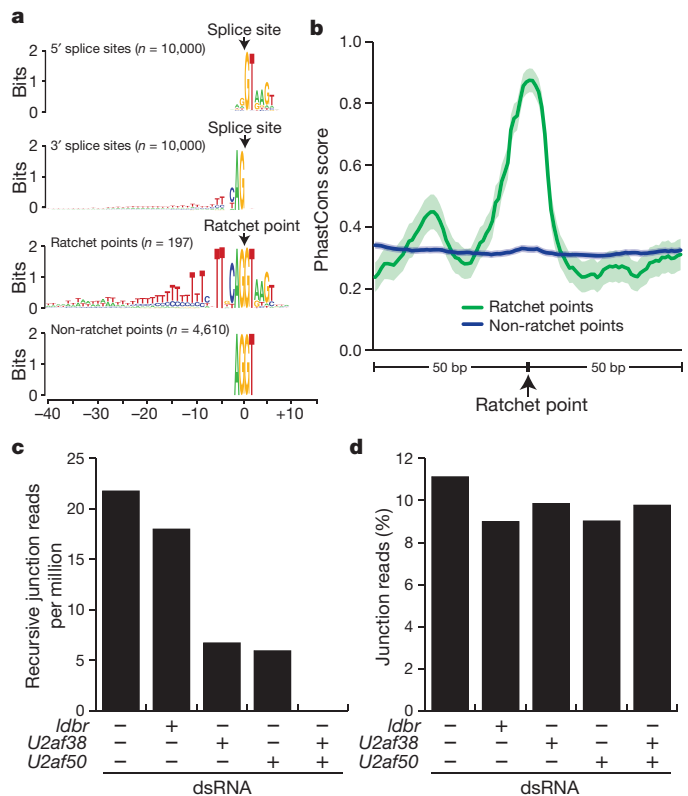


Figure 3 | Characteristics of *Drosophila* ratchet points. **a**, Sequence logos of 5' splice sites, 3' splice sites, ratchet points and non-ratchet point AG/GT sequences located in the same introns as ratchet points (top to bottom). **b**, Sequence conservation of ratchet points. Average PhastCons scores of ratchet points (green) and non-ratchet points (blue). Solid line indicates the average PhastCons score; shaded regions indicate the 95% confidence interval. **c**, **d**, Normalized recursive junction (c) reads and per cent non-recursive junctions (d) observed in untreated S2 cells and cells treated with the indicated dsRNAs.

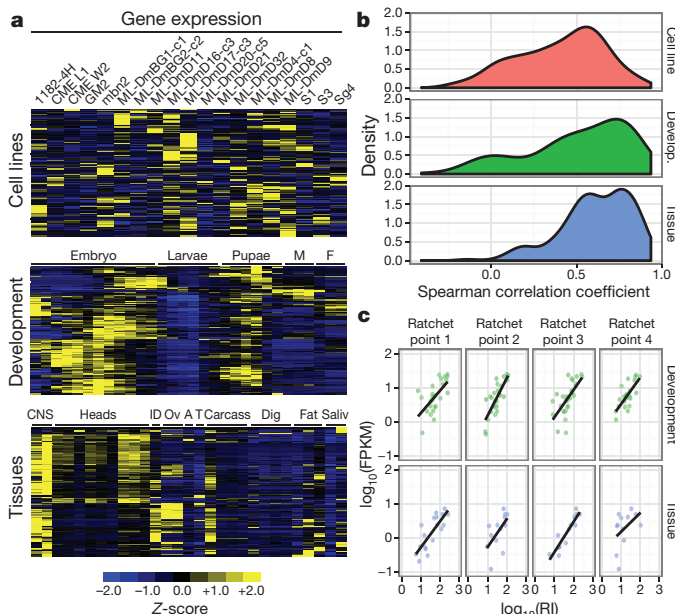


Figure 4 | Expression characteristics of recursively spliced *Drosophila* genes. **a**, Heat-map representation of Z-scores of mRNA expression levels of the recursively spliced genes among the samples examined. A, accessory gland; Dig, digestive tract; F, female; ID, imaginal discs; M, male; Ov, ovaries; Saliv, salivary gland; T, testes. **b**, Distribution of the Spearman correlations of mRNA expression levels and recursive indexes of each ratchet point for the cell line (red), developmental (green), and tissue (blue) samples. **c**, Example of the correlation of mRNA expression levels and recursive indexes (RI) for four ratchet points in *Antp* in the developmental (green) and tissue (blue) samples.

RT-PCR experiments, we failed to amplify lariats generated by ratchet point skipping using primers that successfully amplified lariats from individual recursive segments in the same intron. We also were unable to identify skipping events in the total RNA-seq data, although we did identify lariats in non-recursive introns as large as 84,027 bp in our total RNA-seq data (data not shown). Second, we calculated a recursive index for each ratchet point (the number of ratchet point junction reads/mapped reads) and observed generally strong correlations between the recursive index and the gene expression level for most genes (Fig. 4b). For example, there is a strong positive correlation between gene expression and recursive splicing for all four ratchet points in the *Antennapedia* (*Antp*) gene (Fig. 4c). The correlation between gene expression and recursive splicing is strongest among the tissue samples and weakest among the cell lines, which have the highest and lowest number of mapped reads, respectively (Extended Data Fig. 4), indicating that low correlation is related to sequencing depth. Together, these results strongly suggest that recursive splicing is constitutive, although it remains possible that regulated or alternative ratchet points may be identified in the future.

Recent studies have demonstrated strong associations between chromatin marks and particular features of gene architecture, including intron–exon boundaries. Of particular note, H3K4me3 (ref. 10), H3K79me2 (ref. 11) and H3K36me3 (ref. 12) have been shown to specifically transition near intron–exon boundaries in humans. We inspected ChIP-seq data obtained from whole larvae to determine whether any chromatin marks are associated with ratchet points (Extended Data Fig. 5). None of the chromatin marks we examined is specifically associated with ratchet points, yet the recursive splice sites are associated with chromatin marks that would be expected given their position relative to canonical exons.

Here we provide experimental evidence that 130 *Drosophila* introns, 26 times the number previously known, are removed in multiple, sequential steps by recursive splicing, rather than by a single splicing event. We also identified five ratchet points in four human genes, including one case of orthologous *Drosophila* and human genes, indicating that recursive splicing evolved long ago. We do note, however, that recursive splicing in *Drosophila* involves true zero nucleotide exons but that an accompanying paper⁹ has demonstrated that recursive splicing in humans involves recursive exons, pointing to mechanistic and perhaps functional differences in this process between flies and humans. The ratchet points involved in recursive splicing are highly conserved and share sequence similarity with one another. While recursive splicing clearly occurs in *Drosophila*, its function and mechanism remain elusive, although we provide evidence that U2AF is required for recursive splicing. It also remains unknown why some *Drosophila* introns are recursively spliced and others are not. Further investigation will be necessary to determine whether recursive splicing is required for the function of the host gene in *Drosophila* and how the upstream exons re-engage in subsequent splicing reactions.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 June 2014; accepted 16 April 2015.

Published online 13 May 2015.

1. Hatton, A. R., Subramaniam, V. & Lopez, A. J. Generation of alternative *Ultrathorax* isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions. *Mol. Cell* **2**, 787–796 (1998).
2. Burnette, J. M., Miyamoto-Sato, E., Schaub, M. A., Conklin, J. & Lopez, A. J. Subdivision of large introns in *Drosophila* by recursive splicing at nonexonic elements. *Genetics* **170**, 661–674 (2005).
3. Conklin, J. F., Goldman, A. & Lopez, A. J. Stabilization and analysis of intron lariats *in vivo*. *Methods* **37**, 368–375 (2005).
4. Mackay, T. F. *et al.* The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482**, 173–178 (2012).
5. Graveley, B. R. *et al.* The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**, 473–479 (2011).
6. Brown, J. B. *et al.* Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**, 393–399 (2014).
7. Oesterreich, F. C., Bieberstein, N. & Neugebauer, K. M. Pause locally, splice globally. *Trends Cell Biol.* **21**, 328–335 (2011).
8. Hollins, C., Zorio, D. A., MacMorris, M. & Blumenthal, T. U2AF binding selects for the high conservation of the *C. elegans* 3' splice site. *RNA* **11**, 248–253 (2005).
9. Sibley, C. R. *et al.* Recursive splicing in long vertebrate genes. *Nature* <http://dx.doi.org/10.1038/nature14466> (2015).
10. Bieberstein, N. I., Carrillo Oesterreich, F., Straube, K. & Neugebauer, K. M. First exon length controls active chromatin signatures and transcription. *Cell Rep.* **2**, 62–68 (2012).
11. Huff, J. T., Plocik, A. M., Guthrie, C. & Yamamoto, K. R. Reciprocal intronic and exonic histone modification regions in humans. *Nature Struct. Mol. Biol.* **17**, 1495–1499 (2010).
12. Kolasinska-Zwiercz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genet.* **41**, 376–381 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by NHGRI grant U54HG006994 to S.E.C. (PI) and B.R.G. (co-PI), and R01GM095296 to B.R.G.

Author Contributions B.R.G. and S.E.C. supervised data production. S.O., A.O., M.B. and S.C.G. performed experiments. M.O.D., X.W., A.P., M.B., S.C.G. and B.R.G. performed computational analysis. B.R.G. wrote the paper with input from all authors.

Author Information Sequences are available from <http://www.encodeproject.org> and have been deposited in the Short Read Archive under accession numbers SRP056962, SRP056965 and SRP056969 and those listed in Supplementary Table 1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.R.G. (ggravelev@uchc.edu).

METHODS

No statistical methods were used to predetermine sample size.

RNA collections. The *D. melanogaster* RNA samples used for this study have been described previously^{5,6}. RNA was isolated from *D. simulans*, *D. sechellia*, *D. yakuba*, *D. pseudoobscura* and *D. virilis* mixed adults using TRIzol. Total RNA from 20 human tissues was obtained from Clontech (catalogue no. 636643). RNA from the RNAi experiments was extracted from *Drosophila* S2 cells (Life Technologies, catalogue no. R690-07) treated with 20 µg of dsRNA for 5 days. The S2 cells were not authenticated or tested for mycoplasma contamination. The following dsRNA sequences were used for the RNAi experiments: *ldbr* (AAGCTAGGAGATGCTGAATCTTCTCTTCCAGCAGCAGCAGTGAAGA TGAAGACGAGGAAAGGGAGAAGGTAAAGAAAGCTGCTCCTGTACCTC CACCATCCAAATCTGTTCCCGTGACCAAGTTTCTGGCTCTCGACAAAT GCCTGCCACGTCGTGCTTCTCTGCAAGTGGTAGAGGTACCCAGTGAC CCCATCGAAGGCATCCCCCGCTGGAATACGACGCAGAGTGGCTAGC CATCTTGACAGCTACAAATCAGTTGATTTCAGTGAAGGAGAATATTA TTACCTGCCCGGAAAAAGGCGGGAGAGTTTACAGAGCGATCAAACCT TTACCCCACTGAAGAAGAACTAGAAGCAGTGACCGCAAAGTTTCAG AAACCTCAAGTCCCGGAGAATTTGAGCGCACAGTGCCAGCTTTCGA TCCCGCGGAGCAGTCTGATTATAAGCACATGTTTGTGGATCAACCCA AGGTTCAACTAAACCCCGAGCAATACGTTCTGTGCCACTCTGGGTA TAGACGATC), *U2af38* (AGATGCAAGAACACTACGACAATTTTTTCGAG GACGTGTTTCGTAGAGTGCGAGGACAAGTACGGGGAAATCGAGGAGAT GAACGTGTGCGACAACCTAGGCGACCATCTGGTTCGGCAATGTGTACA TCAAATCCGTAACGAGGCTGATGCGGAAAAGGCGCAACAGATTG AACAAACCGGTGGTTGCGGAGCGGTGTACTCGGAATCTCGCC GGTGACCGACTTCCGCGAGGCTTGTCTGCGGCAGTACGAGATGGGCG AATGTACCGCTCCGCTTCTGCAACTTCATGCATTTGAAGCCCATCT CGCGTAGCTGCGAAGGTACCTCTACTCCCGCCCGCTCGTGCCCGC TCCCGTTCCCGATCCCTGGACGCCGCTCGCGCTCCCGCAGCAGGTCC CGATCCCGGGTTCGAAGAGGAGCGCGCAGAGCGACGCTGTCGGTG GAGGAACTACTTGAACAAC) and *U2af50* (CCGAGGAGGAAATGATG AGTTCTTCAACCAACAGATGCATTTAGTTGGGCTCGCCAGGCGGCC GGCAGTCCCGTCTTGGCATGCCAAATTAAGTTGGACAAAACTTTGCT TTCTCGAATTCGATCGATTGATGAACACCCAGGCCATGGCATT CGATGGCATCAATTTGAAGGGGCGAGCTTAAAGATTAGGCGTCCGC ACGATTACAGCCCATGCCGGGTATAACAGATACGCCGCAATTAAG CCCGCTGTTGTTTCCAGTGGAGTTATTTTCGACAGTGGTTCGGACTCG CCTCACAAAATCTTCATCGGAGGTCTACCAAACTATCTGAATGACGAT CAGGTTAAGGAAGCTTTTGTGCTTGGCAAGCTACGAGCCTTCAAC CTGGTTAAGGATGCCGTACTGGGTTAGTAAGGGTTATGCTTTCTG TGAATATGTCGATCTTAGCATCACAG).

RNA sequencing. Total RNA-seq libraries were prepared using Illumina TruSeq Stranded Total RNA Sample Prep Kits as described by the manufacturer. Libraries were quantified by analysis on an Agilent Bioanalyzer or TapeStation and sequenced on an Illumina HiSeq2000 to generate paired-end 100 bp reads or an Illumina NextSeq500 to generate paired-end 76 bp reads (for the RNAi experiments).

Alignments. Total RNA strand-specific paired-end sequence data from *D. melanogaster* was aligned to the *D. melanogaster* genome (Release 5, dm3) lacking chromosome U extra, guided by the modENCODE annotation MDv1 (ref. 5) using TopHat¹³ version 1.4.1 with the following settings: -p 8 -z 0 -a 6 -m 2-min-intron-length 28 -I 200000 -g 1-library-type fr-firststrand -x 60 -n 2. The *D. simulans*, *D. sechellia*, *D. yakuba*, *D. pseudoobscura* and *D. virilis* RNA-seq data sets were aligned to the *D. simulans* (droSim1), *D. sechellia* (droSec1), *D. yakuba* (droYak2), *D. pseudoobscura* (dp4) and *D. virilis* (droVir3) genomes, respectively, using the same method and parameters, but without a reference transcriptome annotation. The human RNA-seq data sets were aligned to hg19 using the same method and parameters and Gencode v19 as a reference annotation.

Parsing TopHat alignments to identify ratchet points. We identified sets of novel splice junctions from the TopHat alignments which share the same 5' splice site. Ratchet point junctions were kept from a sample if the 3' splice site of the junction is an AG/GT, the 3' splice site was unannotated in the previous modENCODE annotation⁵, and the distance to the previous splice junction and the next splice junction is longer than 2 kb.

De novo identification of ratchet points. We generated a set of potential ratchet point junctions by joining 95 nt of each exon to the 95 nt downstream of every unannotated 3' splice site (AG/GT) in the downstream intron. We aligned reads 1 and 2 of the total RNA reads independently to the database of all possible ratchet point junctions using Bowtie¹⁴ version 0.12.7 with the following options: -v 2 -k 5 -M 5-best. As the paired-end reads were aligned separately, post-processing was

used to enforce constraints on gene-strand and alignment-strand that are a consequence of the stranded protocol. For each potential ratchet-point junction, we tabulated the coordinates of the genomic regions that comprise the ratchet-point junction sequences, the intron(s) and gene(s) the ratchet site is derived from, the number of alignments to the ratchet-point junction, the average number of mismatches per alignment, detailed offset and mismatch information, the alignment offset entropy⁵, and the number of distinct offsets for only perfect alignments with ≥ 8 nt overhang. This latter parameter is intended to be a robust and conservative measure of alignment diversity. Ratchet points contained in introns that overlap no other distinct introns or any exons were filtered to require ≥ 3 perfect alignments to three distinct offsets and overhang ≥ 8 (or two perfect alignments to two distinct offsets and overhang ≥ 8 and ≥ 10 general alignments with ≤ 2 mismatch and overhang ≥ 5). Ratchet points in introns that do overlap other introns or exons were filtered using slightly more restrictive criteria and required at least five zero-mismatch ratchet-junction alignments with distinct offsets.

Verification of potential ratchet points. We next compared the lists of potential ratchet points individually identified by analysis of the TopHat alignments and by alignment to all potential ratchet point junctions. This resulted in a list of 356 potential ratchet points that were then individually examined on the genome browser to verify their identity. To facilitate this analysis we merged the bedGraphs from all of the TopHat alignments into one positive- and one negative-strand-specific bedGraph file. For each intron containing potential ratchet points, we calculated a robust linear regression of the read density of each segment of the recursive introns from the merged bedGraph profile using the robustfit function of MATLAB. This required masking out repeatMasker regions and overlapping annotated features, both of which confound the regression process, and calculating the robust regression lines based on the remaining unmasked portions of the recursive segment. We used MATLAB to generate a 'flip-book' of browser-like images of each potential recursive intron with bedGraph and local robust regression plots superimposed to aid in the manual inspection of each ratchet point for verification.

The merged bedGraphs were loaded into the genome browser along with tracks of all 356 potential ratchet point splice junctions. In addition, we loaded the FlyBase 5.45 annotation, which was the most recent version of FlyBase at the time of this analysis, as well as the most recent modENCODE annotation (MDv3)⁶ to identify ratchet points that corresponded to exons identified more recently than the modENCODE annotation (MDv1)⁵ used to seed the alignments. Finally, we loaded in the modENCODE CAGE data⁵ to assess whether any potential ratchet points corresponded to previously unannotated promoters, which could also give rise to a saw-tooth pattern of RNA-seq read density. Potential ratchet points were removed if the 3' ratchet point junction corresponded to an annotated exon, or if the saw-tooth pattern of read density was not apparent on the browser or from the local robust regression plots. During the course of this manual inspection we identified two ratchet points in *luna* and *mbf* that were not identified in this computational analysis. Both of these were present in introns that contained other computationally identified ratchet points. We identified these based on their strong pattern of saw-tooth read density in both the browser and the local robust regression plots and the fact that they had conserved AG/GT sequences at the ratchet point junctions. These were missed in the computational analysis because they did not pass the stringent filters used. Both of these were experimentally validated and included in the analysis of all ratchet points. In total, the final list of ratchet points consists of 197 ratchet points (Supplementary Table 2). The same approach was used to manually review the putative ratchet points identified in the human RNA-seq data.

Validation of recursive splice sites by RT-PCR. 500 ng of total RNA isolated from S2 cells with TRIzol was used to synthesize cDNA using SuperScript II Reverse Transcriptase (RT) kits according to the manufacturer's protocol. PCR amplification was performed using Phusion High-Fidelity DNA Polymerase (NEB) according to the manufacturer's instructions using specific primers with the following amplification program: 98 °C for 3 min, followed by 40 cycles 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 20 s. Ratchet points were first confirmed by gel electrophoresis followed by Sanger sequencing.

Validation of recursive splice sites by cross-species RNA sequencing. The coordinates of the *D. melanogaster* ratchet points, and 50 nt on either side, were lifted over to the *D. simulans* (droSim1), *D. sechellia* (droSec1), *D. yakuba* (droYak2), *D. pseudoobscura* (dp4) and *D. virilis* (droVir3) genomes using the UCSC liftover tool on galaxy. The TopHat alignments of the *D. simulans*, *D. sechellia*, *D. yakuba*, *D. pseudoobscura* and *D. virilis* were searched for splice junction reads whose 3' end mapped within the lifted over coordinates and which had an AG/GT at the recursive junction.

Identification of recursive lariat introns and branch-point analysis. To generate potential junctions between the 5' splice site and branch points of intron lariats, we used a custom perl script called build_branchpoint_junctions.pl to fuse

the last 94 nt of an intron to the first 94 nt of the intron. We used 94 nt of each portion of the intron to enforce a minimum of a 6-nt overhang when aligning 100-nt reads. Because the precise location of the branch points are not known ahead of time, and because a previous study in human has shown that most known branch points occur between 18 and 35 nt upstream of the 3' splice site¹⁵, we generated 100 possible lariat junctions by sliding a window of the 3' end of the intron in the 5' direction at 1-nt intervals, and fusing each to the first 94 nt of the intron. We generated the potential lariat junction databases for each segment of all recursive introns, as well as all possible permutations of these segments. For example, for an intron with two ratchet points, we generated potential lariat junctions for the first, second and third recursive segments as well as the introns from the first 5' splice site to the second ratchet point (segments 1 and 2), from the 5' splice site of the first ratchet point to the last 3' splice site (segments 2 and 3), and from the first 5' splice site to the last 3' splice site (segments 1, 2 and 3).

Bowtie¹⁴ version 0.12.7 was used to generate an index of the potential recursive lariat junctions and to align all of the total RNA-seq reads, where each mate pair was aligned separately, with the following parameters: -v 2 -p 8—all—quiet. In total, 72,712 alignments were reported. These were then filtered for reads that mapped uniquely to the lariat junction database, then sorted by the number of nucleotides overlapping the lariat junction. Randomly selected reads with various extents of overlap were aligned to the *D. melanogaster* genome using BLAT to determine whether or not they mapped elsewhere in the genome. From this we determined that all reads that overlapped the lariat junction with fewer than 14 nt also aligned elsewhere in the genome. We therefore used BLAT to align all reads reported from the Bowtie alignment and discarded those which mapped elsewhere resulting in a total of 46 reads. We identified the approximate locations of the branch points based on the coordinates of the 94-nt segment of the 3' end of the intron that the reads aligned to.

Validation of recursive lariats. Thirty-eight recursively spliced genes were selected for lariat analysis, choosing genes that were highly expressed in the nervous system. These genes contain 95 distinct segments corresponding to 52 ratchet points. To detect splice lariats for these ratchet points, 'outward-facing' PCR primer pairs were designed to amplify through branch points. The PCR primers contained overhangs so that Illumina clustering, indexing and sequencing oligonucleotides could be added in a subsequent nested PCR. The same primers designed for ratchet point lariat amplification were also used in different combinations to attempt to amplify the branch-point lariats that would be generated by skipping one or more ratchet points.

Total RNA was extracted from whole *D. melanogaster* (Bloomington *Drosophila* Stock Center strain no. 2057) using TRIzol reagent (Invitrogen, Grand Island, NY), followed by cDNA synthesis using Superscript II (Invitrogen) primed with random hexamers. The primer pairs described above were used for the first-round PCR, after which products were visualized on an agarose gel. For 13 genes, no product was detected for any of the sub-introns targeted and these genes were not analysed further. For the other 25 genes, which contained a total of 64 sub-introns, we obtained a PCR product for at least one sub-intron. For these 25 genes we also used the primers in combinations to amplify any potential lariats created if splicing skipped ratchet points in a total of 41 combinations. Regardless of whether or not a product was visualized we prepared sequencing libraries from all reactions in an effort to capture any low-level amplicons. Nested PCR was performed to add the Illumina sequencing oligonucleotides to the first-round PCR products. The amplicon libraries were pooled, purified and size selected to select amplicons between 300 to 1,000 bases in length. The pooled amplicon library was then sequenced on an Illumina MiSeq using a V3, 600-cycle kit to produce 200 by 400 bp paired-end reads.

Reads were filtered to remove mis-primed sequences and then aligned to the target genes using BLAT. Mapped reads were manually reviewed on the genome

browser and a lariat considered to be 'confirmed' if a portion of a read aligned precisely to the 5' splice site (ratchet point or exon) and to a second region within 100 bp upstream of the 3' splice site (ratchet point or exon) in a continuous manner. This analysis confirmed 14 of the 64 recursive segments examined. Importantly, none of the primer combinations used to detect ratchet point skipping yielded any reads corresponding to a lariat. The few sequence reads that were obtained for these controls were PCR artefacts that primarily contained Illumina sequencing oligonucleotides and had no homology to the target genes. Sequence logos were generated with WebLogo¹⁶.

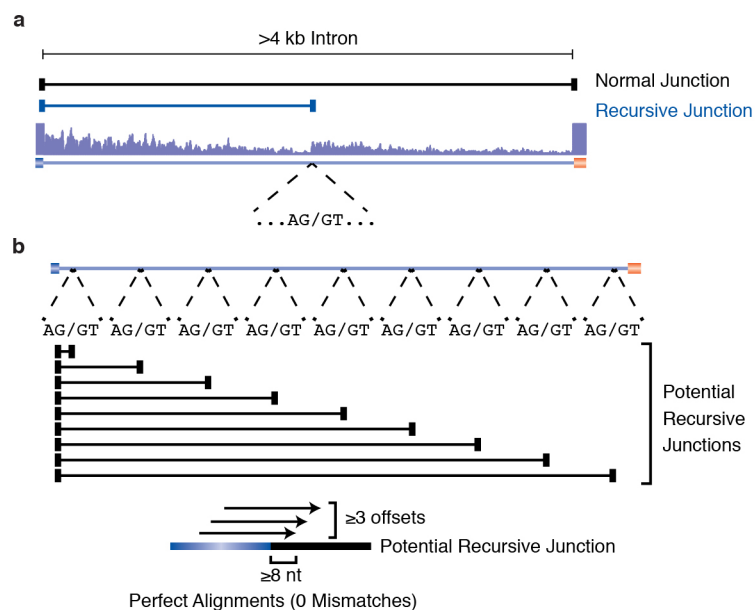
Comparison of gene expression and recursive splicing. The total number of reads mapping to each ratchet point junction was tabulated for each library and then summed across biological and/or technical replicates for each biological sample. The recursive index was calculated as the number of ratchet point junction reads per mapped reads per billion reads ((ratchet point junction reads/mapped reads) \times 1,000,000,000). The RNA-seq data generated for this study is from total RNA, a mixture of mRNA, nascent RNA and pre-mRNA. Owing to the complexity of RNA types and intronic reads it is difficult to accurately quantify gene expression levels when using most existing software. We therefore used the expression values calculated from the corresponding poly(A)⁺ RNA-seq data that was previously generated^{5,6} from the same RNA samples. Comparisons of the recursive index and gene expression levels were performed using R (<http://www.r-project.org>). GO analysis was performed using FuncAssociate 2.0 (ref. 17).

Analysis of chromatin marks. Visualizations of chromatin marks at recursively spliced genes were generated using custom R scripts (<http://www.r-project.org>). We obtained ChIP-seq scores (http://encode-x.med.harvard.edu/data_sets/chromatin/) and Affymetrix tiling array gene expression scores (<http://intermine.modencode.org/>) generated from L3 larvae via the modENCODE projects. For each feature of gene architecture illustrated, mean ChIP-seq scores were calculated for non-overlapping bins of 200 bp in length.

As expected, transcription-associated marks were specific to genes actively transcribed in larvae (Supplementary Fig. 5b). At these active genes, we observed low levels of H3K4me3 near recursive splice sites compared to first exons (Supplementary Fig. 5b), which suggests that the saw-tooth patterns observed by total RNA-seq were not due to cryptic transcription initiation or unannotated promoters, but rather, co-transcriptional splicing. We also observed lower levels of H3K36me3 near recursive splice sites compared to downstream exons (Supplementary Fig. 5b). Since the degree of H3K36me3 has been shown to increase with each internal exon in humans¹⁸, the low levels of H3K36me3 seen at recursive splice sites may reflect the fact that recursive splices are typically located in 5' introns, and thus, preceded by few internal exons (Supplementary Fig. 5c). Indeed, recursive splice sites were associated with high levels of H3K79me2 exons, which is typical of long 5' introns in humans¹¹.

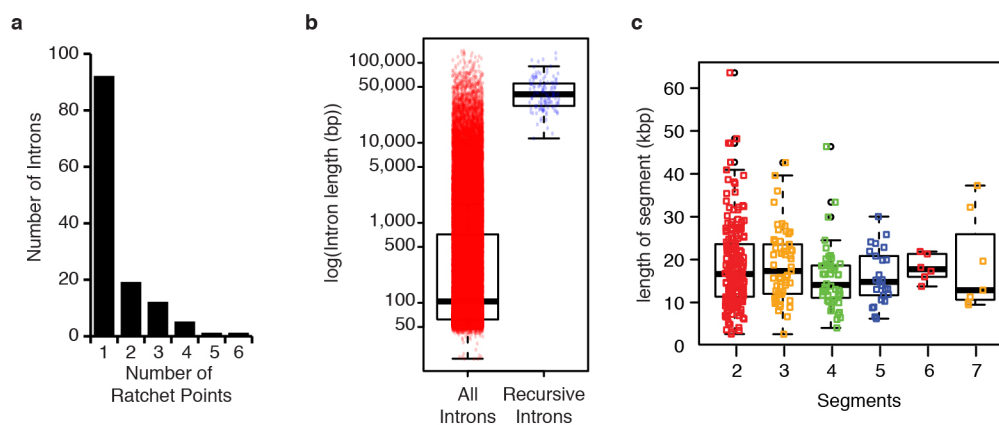
Code availability. Custom code used in this paper is available without restrictions at <https://github.com/graveley/Recursive>.

13. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
14. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
15. Taggart, A. J., DeSimone, A. M., Shih, J. S., Filloux, M. E. & Fairbrother, W. G. Large-scale mapping of branchpoints in human pre-mRNA transcripts *in vivo*. *Nature Struct. Mol. Biol.* **19**, 719–721 (2012).
16. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
17. Berriz, G. F., Beaver, J. E., Cenik, C., Tasan, M. & Roth, F. P. Next generation software for functional trend analysis. *Bioinformatics* **25**, 3043–3044 (2009).
18. Tilgner, H. *et al.* Nucleosome positioning as a determinant of exon recognition. *Nature Struct. Mol. Biol.* **16**, 996–1001 (2009).



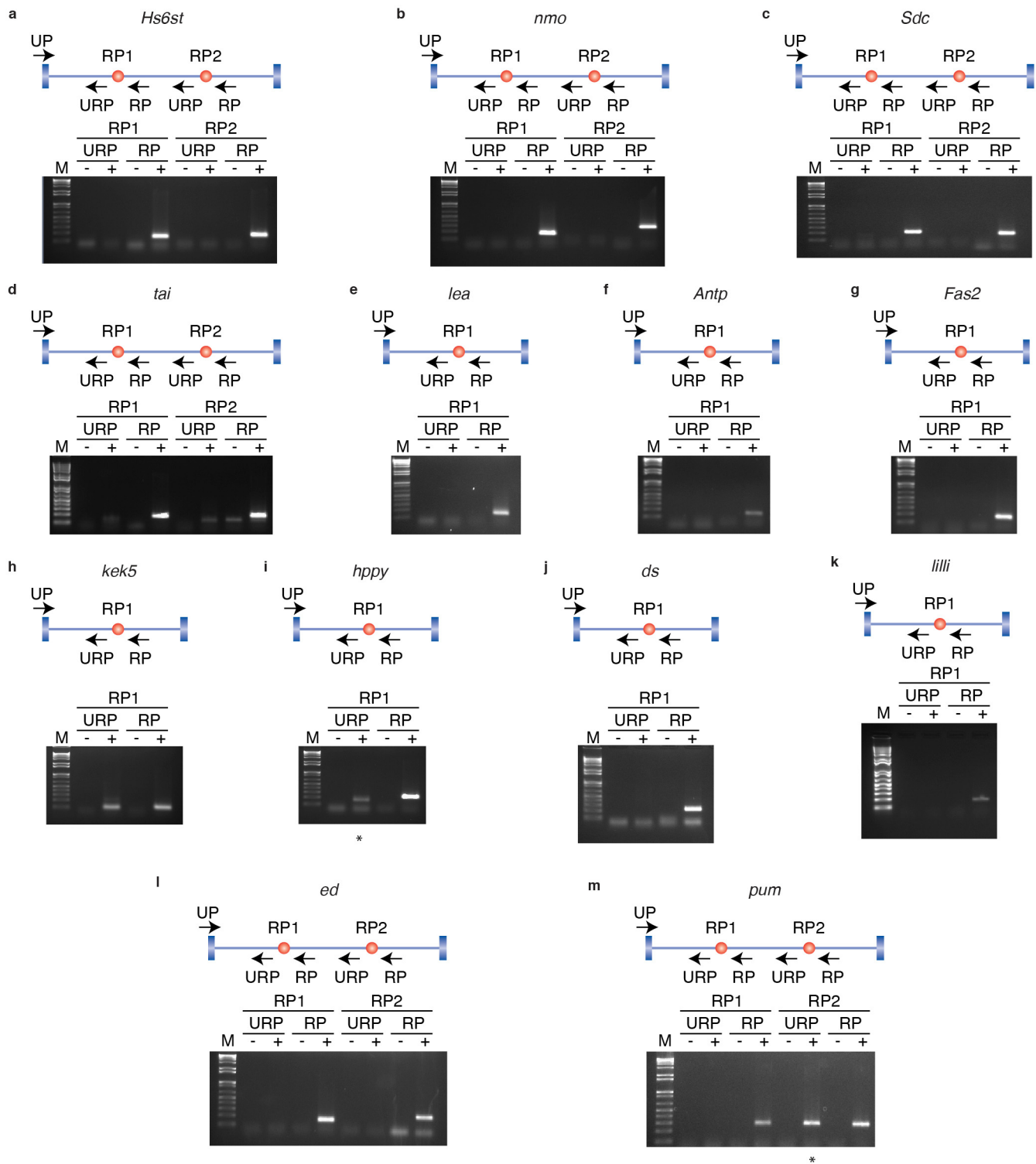
Extended Data Figure 1 | Two approaches for identifying recursive splice sites. **a**, Identification of recursive splice sites by parsing alignments. RNA-seq reads were mapped to the genome using TopHat in a manner that allowed for novel splice junctions to be predicted. The alignments were then parsed for splice junction reads where the 5' splice site mapped to an annotated 5' splice site, but the 3' splice site was unannotated. **b**, *De novo* identification of recursive

splice sites. A database was generated in which each annotated 5' splice site was spliced to all AG/GT sequences in an intron that did not correspond to an annotated 3' splice site. All RNA-seq reads were aligned to this database and the alignments parsed to find cases where reads mapped perfectly with at least three distinct offsets and at least an 8 nt overhang.



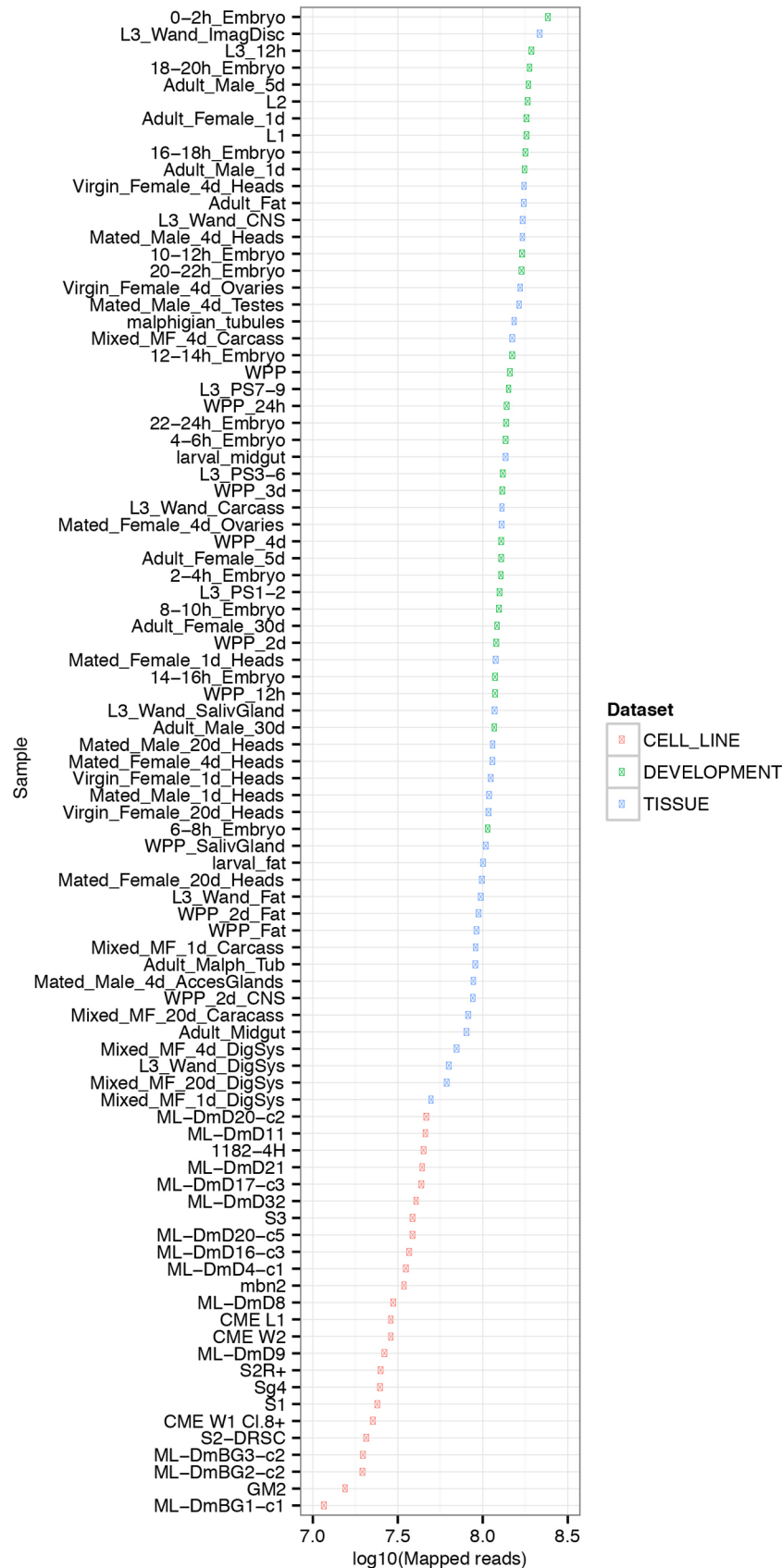
Extended Data Figure 2 | Characteristics of *Drosophila* ratchet points.
a, Distribution of the number of ratchet points per recursive intron. **b**, Size distribution ($\log_{10}(\text{bp})$) of all (red) and recursive (blue) introns. **c**, Size

distribution (in kb) of the individual intron segments removed by recursive splicing binned by the number of segments per intron.

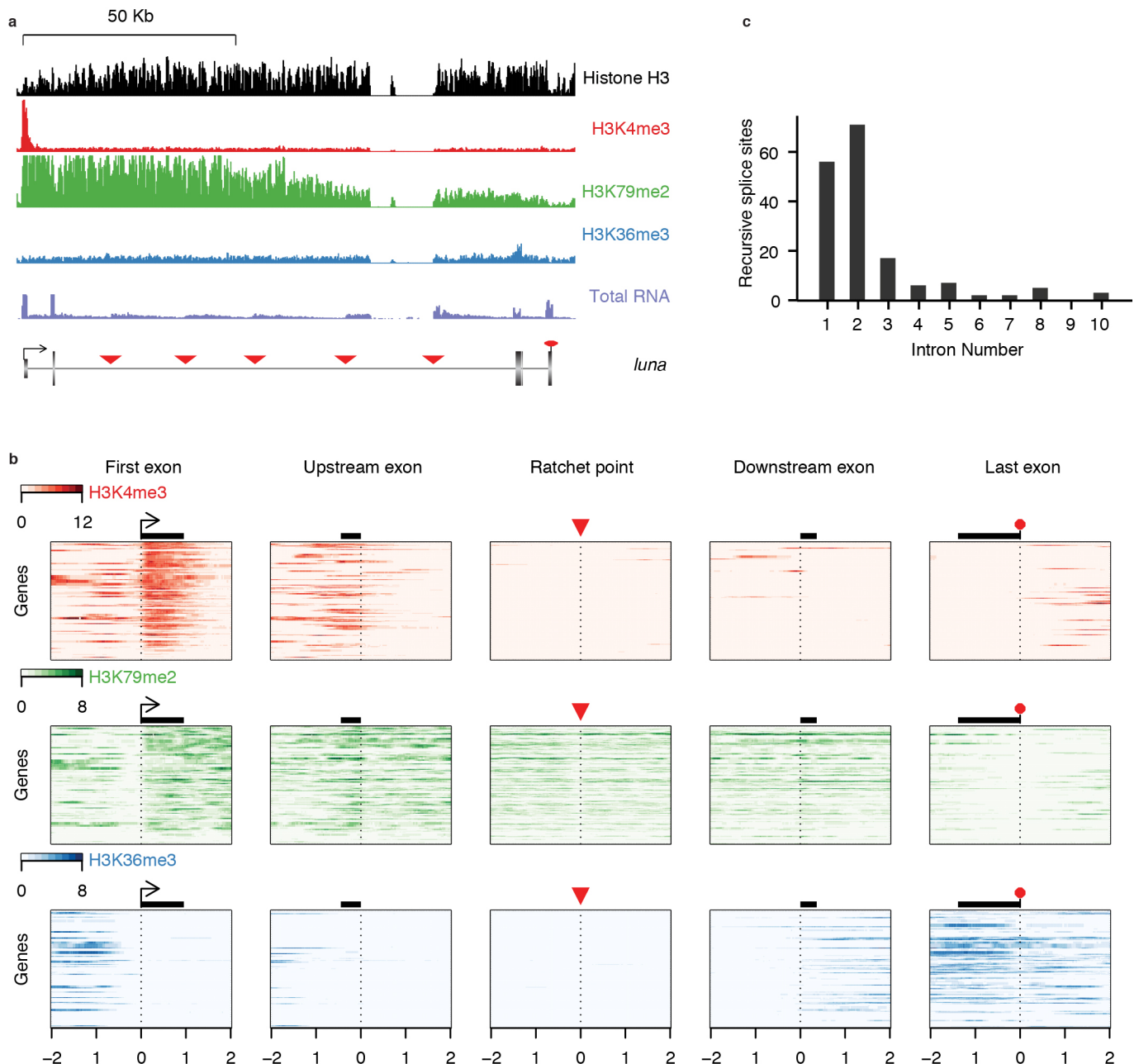


Extended Data Figure 3 | RT-PCR validations of *Drosophila* recursive splicing events. **a–m**, RT-PCR validation of ratchet points (red dots) from the indicated genes using primers in the upstream constitutive exon and flanking the putative ratchet points. The RP primers are expected to yield RT-PCR products if the constitutive exon is spliced to the ratchet point. The URP

primers, which are upstream of each ratchet point, serve as negative controls. The identity of all RT-PCR products was verified by Sanger sequencing. Although the URP control RT-PCR reactions yielded a product for *hppy* RP1 and *pum* RP2, we were not able to generate sequence from them and therefore consider them to be amplification artefacts.



Extended Data Figure 4 | Number of mapped reads per sample used for gene expression analysis.



Extended Data Figure 5 | Chromatin marks associated with recursive splice sites. **a**, Examples of chromatin marks at the *luna* gene locus, which contains 5 recursive splice sites (red triangles) within a single long intron. **b**, Heat maps show relative ChIP-seq enrichment for H3K4me3 (top, red), H3K79me2 (middle, green) and H3K36me3 (bottom, blue), within 2 kb of the indicated gene features from 171 genes containing at least one ratchet point. Heat maps are centred around gene features, which include the transcription start site of the first exon (first exon, arrow), the 5' splice site of the exon upstream of the

recursive splice site (upstream exon, black rectangle), the ratchet point (red triangle), the 3' splice site of the exon downstream of the recursive splice site (downstream exon, black rectangle), and the poly(A) site of the last exon (last exon, red octagon); the average exon of each gene feature is drawn to scale. Genes are sorted from top to bottom by decreasing expression level. For genes containing more than one ratchet point, the first, upstream, downstream and last exons are represented multiple times. **c**, Histogram illustrating the intron positions the ratchet points reside in based on RefSeq annotations.

Extended Data Table 1 | Summary of recursive intron lariats identified by directed RT-PCR and sequencing

Gene	Segment	Coordinate of putative branchpoint	Distance upstream of 3' splice site	Sequence surrounding bp	Identified in Total RNA-Seq Data
<i>Sdc</i>	segment3	2R:17299944	21	CATCTCACTCATAAATGTGTT	No
<i>cpo</i>	segment1	3R:13803044	34	AAGGTAACATAATGATTTTT	Yes
<i>cpo</i>	segment2	3R:13815266	31	CCAAATGCTAATTTTATACTT	Yes
<i>cpo</i>	segment3	3R:13832619	37	AGCAATCATCTAACGATTCTC	Yes
<i>bun</i>	segment2	2L:12458256	32	CAACATACTTACAGAACCTTT	No
<i>CG7029</i>	segment2	3R:18592273	55	GTTTGTGCTCACAGAGTCTGC	No
<i>nuf</i>	segment2	3L:14223246	29	ATATAGACTTATCAGTTCTCT	No
<i>CG31637</i>	segment2	2L:6525343	23	GAGTATTCTAACAAGTTTCTC	Yes
<i>dally</i>	segment1	3L:8843654	26	CTAAATCTGTGCTTAATTTCT	No
<i>dally</i>	segment2	3L:8855584	45	AATTTGCACCATCGCATAACT	Yes
<i>dally</i>	segment3	3L:8870223	29	ATCCAAGCTCATCTCCTCTTT	Yes
<i>Mmp2</i>	segment2	2R:5503040	33	TAGCATGCTGATATCATGTTT	No
<i>osp</i>	segment2	2L: 14656399	30	AACCAAACTAATTTTTCTACC	No
<i>osp</i>	segment1	2L:14677529	41	ACATCTTCTTACTAAATTATT	No

For each recursive lariat confirmed by RT-PCR and sequenced, the following information is provided: gene, segment, coordinate of the putative branch point, distance upstream of the 3' splice the branch point is located, the sequence surrounding the branch point, whether the lariat intron is newly validated with respect to the total RNA analysis described in Supplementary Table 4a.

Extended Data Table 2 | Summary of total RNA-seq data from *Idbr* and *U2AF* RNAi experiments

Sample	S2 Untreated	<i>Idbr</i> dsRNA	<i>U2af38</i> dsRNA	<i>U2af50</i> dsRNA	<i>U2af38</i> & <i>U2af50</i> dsRNA
Total reads	183,451,394	146,237,470	174,626,770	162,012,182	179,417,376
Mapped reads	66,597,777	54,762,604	56,120,630	43,620,982	61,437,507
% mapped	36.30%	37.45%	32.14%	26.92%	34.24%
# junctions	44,860	42,736	52,228	46,184	50,883
junctions/1M reads	673.60	780.39	930.64	1058.76	828.21
# total junc reads	7,402,459	4,925,448	5,528,462	3,935,048	5,999,468
% junction reads	11.12%	8.99%	9.85%	9.02%	9.77%
# novel junction reads	2,369,949	1,575,057	1,337,830	919,616	1,273,867
% novel junction reads	3.56%	2.88%	2.38%	2.11%	2.07%
# annotated junc reads	5,032,510	3,350,391	4,190,632	3,015,432	4,725,601
% annotated junction reads	7.56%	6.12%	7.47%	6.91%	7.69%
#novel junction cases	22,688	20,977	29,211	24,279	28,193
novel/total junctions	0.51	0.49	0.56	0.53	0.55
# annotated junc cases	22,172	21,759	23,017	21,905	22,690
annotated/total junctions	0.49	0.51	0.44	0.47	0.45
#RPs on both strands	119	100	81	64	0
RP/million reads	1.79	1.83	1.44	1.47	0.00
Total RP junc reads	1449	985	377	259	0
RP junc reads/million	21.76	17.99	6.72	5.94	0.00

For each data set the following information is included: total reads, mapped reads (from TopHat alignments), % mapped, number of total splice junctions, normalized junctions, total number of splice junction reads, per cent splice junction reads, number of novel junction reads, per cent novel junction reads, number of annotated junction reads, per cent annotated junction reads, number of novel junction cases, ratio of novel to total junctions, number of annotated junction cases, ratio of annotated to total junctions, number of ratchet points, normalized ratchet points (ratchet points per million reads), total ratchet point junction reads, ratchet point junction reads per million reads.

Extended Data Table 3 | Summary of total RNA-seq data from related *Drosophila* species

Species	Total reads	Mapped reads	% mapped
<i>D. simulans</i>	92,949,544	45,797,569	49.27%
<i>D. sechellia</i>	38,674,020	19,974,783	51.65%
<i>D. yakuba</i>	46,002,798	18,300,604	39.78%
<i>D. pseudoobscura</i>	45,274,524	14,457,616	31.93%
<i>D. virilis</i>	46,675,210	18,544,567	39.73%
Total	269,576,096	117,075,139	

For each data set the following information is included: species, total reads, mapped reads (from TopHat alignments), % mapped.

Extended Data Table 4 | Summary of ratchet points experimentally validated in other *Drosophila* species

Species	Total leftovers	number of ratchet junction reads	distinct ratchet points	% validated
<i>D. sechellia</i>	163	93	74	45.40%
<i>D. simulans</i>	166	106	87	52.41%
<i>D. yakuba</i>	150	52	43	28.67%
<i>D. pseudoobscura</i>	78	15	15	19.23%
<i>D. virilis</i>	40	18	18	45.00%

For each species analysed, the following information is provided: species, total leftovers (the number of *D. melanogaster* ratchet point coordinates that were successfully lifted over to the other species genome coordinates), number of ratchet junction reads (the number of reads that mapped to the AG/GT sequence of the lifted over ratchet points), distinct ratchets (the number of distinct ratchet points identified by the mapped reads), % validated (the per cent of lifted over ratchet points with at least one mapped read).

Extended Data Table 5 | Summary of total RNA-seq data from human tissues

Tissue	Total Reads	Total Aligned Reads	Unique Aligned Reads	% Uniquely Aligned
adrenal-gland	75,843,326	66,482,852	59,958,580	79.06%
brain-cerebellum	70,868,098	64,382,088	60,087,066	84.79%
brain-whole	75,684,718	63,391,490	59,015,902	77.98%
fetal-brain	77,532,892	65,028,630	58,748,974	75.77%
fetal-liver	69,092,806	47,109,928	38,162,898	55.23%
heart	79,200,464	70,490,860	65,388,604	82.56%
kidney	65,109,248	57,411,382	52,917,602	81.28%
liver	67,086,550	58,934,448	50,433,480	75.18%
lung	66,115,012	58,503,662	51,846,946	78.42%
placenta	70,985,882	62,429,622	56,803,382	80.02%
prostate	68,767,818	62,197,480	56,967,700	82.84%
salivary-gland	77,811,154	73,634,924	55,954,594	71.91%
skeletal-muscle	75,606,186	63,492,280	59,413,946	78.58%
small-intestine	72,492,290	62,291,298	57,178,572	78.88%
spleen	82,675,794	62,553,250	55,318,774	66.91%
stomach	73,018,660	62,932,016	56,174,974	76.93%
thymus	72,695,752	67,467,912	60,657,380	83.44%
thyroid	75,794,702	68,200,264	63,259,560	83.46%
trachea	70,890,274	58,922,766	53,022,564	74.80%
uterus	72,876,872	61,974,024	56,835,240	77.99%
total	1,460,148,498	1,257,831,176	1,128,146,738	

For each data set the following information is included: sample, total reads, total aligned reads, uniquely aligned reads, % uniquely aligned.

CORRIGENDUM

doi:10.1038/nature14371

Corrigendum: Three keys to the radiation of angiosperms into freezing environments

Amy E. Zanne, David C. Tank, William K. Cornwell, Jonathan M. Eastman, Stephen A. Smith, Richard G. FitzJohn, Daniel J. McGlinn, Brian C. O'Meara, Angela T. Moles, Peter B. Reich, Dana L. Royer, Douglas E. Soltis, Peter F. Stevens, Mark Westoby, Ian J. Wright, Lonnie Aarssen, Robert I. Bertin, Andre Calaminus, Rafaël Govaerts, Frank Hemmings, Michelle R. Leishman, Jacek Oleksyn, Pamela S. Soltis, Nathan G. Swenson, Laura Warman & Jeremy M. Beaulieu

Nature **506**, 89–92 (2014); doi:10.1038/nature12872
corrigendum *Nature* **514**, 394 (2014); doi:10.1038/nature13842

Three readers pointed out that in this Letter we applied the threshold of 0.044 (the size at which freezing-induced embolisms are believed to become frequent at modest tensions) to the area of the conduit (in mm²) rather than the diameter (in mm). As a consequence, our analysis assumed far too few extant taxa as having a large conduit diameter, which altered the quantitative results considerably for conduit diameter. We now show that (1) the state combination with the largest persistence time is 'large' conduit 'freezing unexposed'; (2) there are fewer transitions out of 'large' conduit 'freezing exposed' than we previously reported owing to many more extant taxa exhibiting this particular state combination; and (3) climate occupancy is more labile than conduit diameter (that is, the ratio of climate to trait is 5.67). Although these quantitative results change for conduit diameter, the interpretation of the possible pathways from 'large' conduit 'freezing unexposed' to 'small' conduit 'freezing exposed' is still qualitatively

the same at the 0.044-mm-diameter threshold. That is, we still find the trait is more likely to evolve prior to a shift in climate occupancy (the trait-first interpretation) at 53.5%. The trait-first pathway, however, is no longer supported for the secondary 0.030-mm-diameter threshold reported on page 11 under "Coordinated evolution of growth habit, leaf phenology, and conduit diameter with climate occupancy" of the Supplementary Information of the original Letter.

The original Letter has not been corrected online. Figure 1 of this Corrigendum shows the corrected Fig. 2b and d. The Supplementary Information of this Corrigendum shows the corrected Extended Data Tables 2, 3 and 4 of the original Letter, with updated conduit diameter results in Extended Data Tables 2 and 3 and updated $-\ln L$ for the AABCD model in Extended Data Table 4. Please refer to the corresponding author A.E.Z. for additional details. We thank E. Edwards, J. deVos and M. Donoghue for bringing this issue to our attention.

Supplementary Information is available in the online version of this corrigendum.

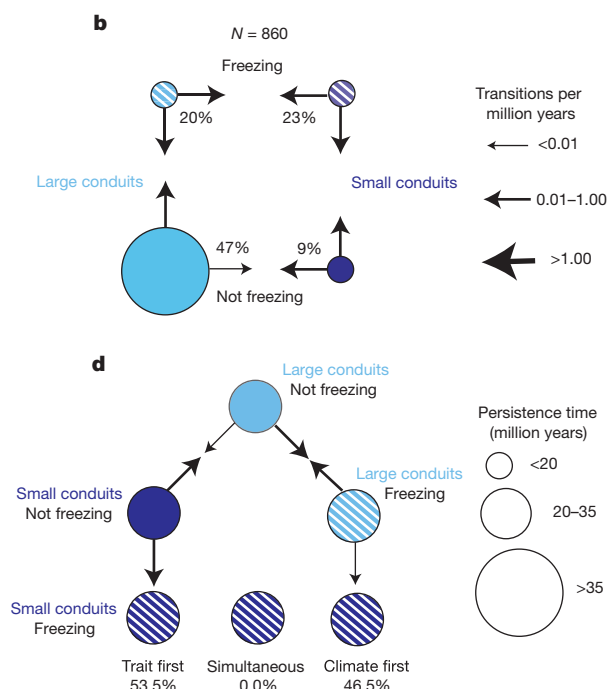


Figure 1 | This is the corrected Fig. 2b and d of the original Letter.

CORRIGENDUM

doi:10.1038/nature14487

Corrigendum: Oxidant stress evoked by pacemaking in dopaminergic neurons is attenuated by DJ-1

Jaime N. Guzman, Javier Sanchez-Padilla, David Wokosin, Jyothisri Kondapalli, Ema Ilijic, Paul T. Schumacker & D. James Surmeier

Nature **468**, 696–700 (2010); doi:10.1038/nature09536

In Fig. 2a of this Letter, the neuron reconstruction and the electrophysiology/calcium imaging traces were mismatched. In addition, in Fig. 2c, the wild-type SNc neurons should have been ' $(n = 5)$ ' rather than ' $(n = 9)$ '. This was a typographical error, and the figure and error bars remain correct. These have all now been corrected in the online versions of the manuscript.

CAREERS

CARDS FOR SCIENCE The astronomer behind a popular parlour game **p.383**

SCIENTIFIC CULTURE Work-life balance needs a rethink **go.nature.com/kohiba**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

MIGUEL NAVARRO/GETTY



RELOCATION

Out of place

Enforced mingling and straight-up instruction can help scientists in a foreign country.

BY PAUL SMAGLIK

When Juhn-Jong Lin relocated from Taiwan to Indiana in 1981, he had few problems picking up the culture of his new laboratory — but social events were another thing entirely. Conversations often centred on pop-culture icons, sports events or political figures — most of which were unfamiliar to him. The patience and goodwill of his labmates made his time at Purdue University in West Lafayette pleasant and productive, but he also experienced some awkward moments, such as when he went to see a classical play with a colleague. “I could not understand it at all,” he says. He remembers that the actors had strong

accents — and he struggled to follow the plot.

Science is an international language, and the assumption behind that cliché is that life beyond the bench will fall into place no matter where the scientist is working. But just because a scientist is fluent in next-generation sequencing does not mean that he or she will know the cultural protocols. For example, interrupting a lab talk with questions is common practice in the United States; in Germany, it is considered rude. And when researchers feel uncomfortable interacting with their colleagues, they will have a hard time doing their best work.

The international nature of science means that many researchers find themselves in a culturally alien situation at some point in

their career. People who plan to work or study overseas can take steps to acclimatize: they can learn beforehand about any introductory programmes that their new institution has set up, and make sure to participate once they arrive. Socializing informally, taking language lessons and creating support networks inside and outside the laboratory can help researchers to navigate unfamiliar, unspoken norms and ease into a foreign culture.

Spurred by growing numbers of young scientists who arrive from other nations as postdoctoral researchers or graduate students, institutions, investigators and administrators are increasingly aware that they need to put out cultural signposts for foreign nationals. Many now provide formal and ►

CULTURE CLASS

Learning how to fit in

Some organizations have created programmes to ease the way for early-career researchers coming from other nations. The US National Institutes of Health (NIH) in Bethesda, Maryland, boasts some 3,500 postdoctoral research and clinical fellows. Of these, 60% hail from other countries, says Sharon Milgram, director of the agency's Office of Intramural Training and Education. Because of the size and complexity of that mix, the NIH puts out a series of programmatic welcome mats.

Visiting fellows can take a two-day course called 'Improving Spoken English', which looks at US cultural norms as well as verbal communication, covering topics such as gun culture, gay marriage, abortion and racism. And natives and newcomers alike can take workshops on workplace dynamics.

Milgram explains that the courses take account of cultural differences; for example, she has employees from cultures in which giving supervisors critical feedback is frowned on, especially in large group settings, so she solicits feedback by e-mail. Pedro Milanez-Almeida, a postdoctoral researcher in immunology at the NIH, says the formal courses made the transition from Germany, where he did his PhD, much smoother than the move from his native Brazil to Germany. "In Brazil, 'tomorrow' means maybe next week," he says. In both Germany and the United States deadlines are strict, but US supervisors expect trainees to take part in setting them.

Early orientation is key, says Ramesh Pillai, a group leader at the European Molecular Biology Laboratory (EMBL) in Grenoble, France. Students and postdocs in a foreign country are often overwhelmed by the basic logistics of life when they first arrive, he says. "Where do you get your food? Where do you buy your bus tickets?"

EMBL uses bureaucracy to help newcomers to create a support network. Fresh arrivals receive a list of contacts for various services, and must get a signature from each person before they receive their meal card or Internet connection. Trainees are also paired with a mentor, and group leaders meet each pair every week. That early, structured interaction prevents social isolation. "The first few days are when you make an impression of a place," Pillai says. "You feel happy about a place, or not." **P.S.**



Ritwick Sawarkar (left) and members of his laboratory, who came to Germany from many countries.

► informal ways to make guests feel more comfortable (see 'Culture class').

There are a range of things that hosts can do to create a better experience for their recruits, say several scientists who were once 'fish out of water' but now manage expatriates of their own. Labs can help to alleviate a feeling of isolation by sponsoring social events. And institutions can prevent some awkward moments by preparing foreign nationals for their host country's cultural norms. Universities that provide logistical support, such as banking or transportation, help new arrivals to weather that first rough week. Places that make a foreigner feel at home in their new country — not just in the lab — are likely to nurture the happiest scientists.

CULTURAL PIONEER

The most empathetic managers may be those who once had to find their footing in an alien land themselves. Ritwick Sawarkar remembers when he moved from India to do a postdoc in Switzerland and has used the experience to shape how he runs his cell-biology lab at the Max Planck Institute in Freiburg, Germany. "I try to make sure that people speak English in my lab as much as possible," says Sawarkar, who has trainees from the United States, Asia and Europe. He wants everyone to have a common language for work and off-work hours. "I've seen what it's like when everyone's speaking German around you. You feel a little bit left out."

Neurobiologist Martin Giurfa can relate. When Giurfa moved from Argentina to Berlin in 1991 to pursue a postdoc, he found the relative solitude stifling. He was used to chatting and socializing with labmates and students, but that was not what he found in Berlin. "You spend the days isolated in a lab, not speaking to anyone," he says. Six months of learning German — paid for by his fellowship — could not bridge the cultural

differences. In his home country, colleagues consider one another to be close personal friends, even if they have just started to work together. The Berlin lab was not like that. "The people were not necessarily there to be friends with each other," he says.

Lack of contact coupled with the pressure to produce results made for a miserable first year. It also did not help that Giurfa's fellowship stipend was pegged to the former East German economy, so his income was scanty as inflation set in after the Berlin Wall fell. But scrambling for cash may actually have salvaged his postdoc experience. He began playing guitar at local nightclubs, which helped with his isolation, his German, his cash flow — and his mindset.

He went on to publish papers, win grants, form collaborations and forge friendships. He is now working at France's basic-research agency (CNRS) in Toulouse. "I am extremely thankful," he says. "Germany made me, in a way."

Giurfa thinks that the transition would have been easier had he known what to expect of his new colleagues. Sawarkar agrees: the cell biologist found that his labmates in Switzerland did not respond positively to his natural effusiveness. "I am usually chatty," he says. "I ask a lot of questions." A friend finally pointed out that his enthusiasm was being interpreted as aggressiveness. Sawarkar took some comfort from learning that he was not alone. "A couple of my American friends told me they had the same experience."

Such shared experiences between outsiders can do much to relieve culture shock and create a social network both inside and outside the lab. These 'safety nets' can then help an expatriate to avoid or deal with a faux pas.

That is what smoothed the transition for Ramesh Pillai from India to Switzerland. His adviser recommended that he stay in a student hostel, rather than rent an apartment for

themselves. Pillai, who is now a group leader at the European Molecular Biology Laboratory in Grenoble, France, says that was some of the best advice he ever received.

Hostel residents help each other to learn the local language as well as where to shop, bank and do laundry. His only misstep was that he focused on learning French so that he could chat with a romantic interest at the hostel, rather than German, the language of informal conversation outside the lab.

Pillai's four-year residence at the hostel also created a social circle for him beyond the laboratory. People from many nationalities mingled in the kitchen nightly to exchange recipes, horror stories and advice, and there were parties almost every weekend. "It was a great atmosphere for anyone coming into a foreign country to meet up with people in similar situations," Pillai says. He tries to create a similar convivial atmosphere in his lab.

BUILDING COHESION

Creating a sense of belonging for lab members leads to greater cohesion and a more effective laboratory. Giurfa has found that young scientists from some countries tend to be overly formal and deferential. "I want them challenging my views," he says. "This positive confrontation could enrich our work more than just agreement." Once lab members feel comfortable with each other, they can communicate more freely. "People are more engaged and productive."

Giurfa also tries to recreate the friendly atmosphere of his native country in his lab team by taking students to vineyards, the Pyrenees and local chateaux. And food provides one of the best ice-breakers for his lab staff: he holds periodic potluck parties, in which lab members bring a dish from their home country and explain why and when it would be served in their culture. In this way, cultural differences form the basis of a shared activity.

Lin remembers how pleased he was when labmates in the US Midwest made the effort to invite him out for drinks. Now a physicist at National Chiao Tung University in Taiwan, he tries to create a hospitable environment for visiting scientists. He says that the world has become more global since his days in the United States: most US and European visitors have already mastered chopsticks and know their way around all manner of Asian cuisine. Many have made an effort to read a bit about Taiwanese politics and culture, or at least read a few articles on Wikipedia. And if conversation stalls, he is ready with his own supply of stories of being a young scientist in an unfamiliar country. ■

Paul Smaglik is a freelance writer in Milwaukee, Wisconsin.

TURNING POINT

Josh Dillon

Astrophysicist Josh Dillon is finishing his PhD at Massachusetts Institute of Technology in Cambridge in an emerging field of cosmology. He is also co-creator of the bawdy card game 'Cards Against Humanity', which this year produced an add-on deck of 30 science-based cards, profits from which will fund a scholarship for women in science.

What does your PhD research involve?

I am working in a field called 21-centimetre cosmology. We're trying to get a baby picture of the Universe. We want to measure the characteristics of the Universe from the time when its first galaxies were forming, about a billion years ago. To do this, we use telescope arrays to detect 21-cm radio waves that were emitted by hydrogen atoms, which were abundant between galaxies then. The challenge is that if the signal exists, it's very faint and is obscured by much more powerful signals from galaxies.

Does this field require new telescopes?

Yes. I've worked on the proposal for a telescope array called HERA (the Hydrogen Epoch of Reionization Array), a huge hexagonal grid of dishes to be built in the Karoo desert of South Africa. We've been using the Murchison Wide-field Array in Western Australia. HERA will be bigger by a factor of about 20, and therefore much more sensitive. These types of array need to be in radio-quiet, remote places; we are monitoring frequencies of 100–200 megahertz, so we want to mitigate interference from FM radio stations that transmit at around 100 megahertz.

Is it scary to work in an unproven field?

Yes and no. I'm pretty optimistic about the field. It has enormous potential. In the 2010 decadal survey of astronomy and astrophysics conducted by the US National Academy of Sciences (go.nature.com/i3v1qj), HERA was one of the highest-ranked projects for ground-based astronomy. It's risky and may not work out as well as we would like. Our biggest challenge is that we may not be able to detect those radio emissions — but all scientific endeavours have risk, and I'm convinced that 21-cm cosmology is worth the risk given the scientific potential. I'm headed this autumn to the University of California, Berkeley, for a postdoc and will work with the team I've been competing against to find the signal.

How does Cards Against Humanity fit in?

It's a fun and worthwhile side project outside my astrophysics pursuits. It started when seven of my high-school friends and I played a card



game that we made up at a New Year's Eve party. It is a politically incorrect party game in which players compete to make the funniest combination of cards. In 2010, we launched a Kickstarter campaign to fund the first print run.

How will the scholarship work?

We formed a board of 40 female scientists to judge a competition to find a candidate who is not only a promising researcher but can also communicate effectively to the public about what she does. We plan to host videos or blog posts to showcase what the winner is doing.

What has the response been like?

Overwhelmingly positive. To date, we've raised more than US\$374,000, so we'll be able to fund at least one or two women. Hopefully, we'll be able to fund more in years to come, depending on how much we raise and the outlay per student. Funding just one scholarship doesn't move the needle that much, but that's only part of why we're doing this. The whole point is to raise the visibility of women in science.

What prompted you to create the scholarship?

Cards Against Humanity has backed other charities, including Wikipedia, Donors Choose, which funds teachers who are eager to do a classroom project, and the Sunlight Foundation, which promotes transparency in politics. When we decided to do a science-focused add-on deck, we knew that we would give the sales proceeds to a charity. We decided that a scholarship for women pursuing an advanced science degree was really appealing. As a company that makes a bawdy party game with a broad social-media reach, we can do one thing — we can help to change the perception of who can be a scientist. ■

INTERVIEW BY VIRGINIA GEWIN

GRAINS OF WHEAT

A lesson learned.

BY ALEX SHVARTSMAN

As he lay dying, Bryce Green contemplated the irony of his predicament. He'd spent a lifetime building the world's foremost pharmaceuticals company. Under his leadership, Green Industries had eradicated numerous ailments and made him the world's seventh-richest man in the process.

The genetic disease ravaging his body was so rare that it had never made financial sense to look for the cure. And by the time he'd learned that it afflicted him personally, it was far too late. His researchers worked feverishly, yet the breakthrough was months, perhaps years away. The doctors told him he had only a few days left.

"There is a woman asking to see you," said his assistant. "She's Rajan Jethwani's daughter."

All sorts of people sought an audience; bootlickers and sycophants, hoping to remind Bryce of their existence, in case there was somehow room for them in his will. He tolerated precious few visitors, and certainly not the child of a one-time business partner from decades ago.

He tried to wave his arm in dismissal, an IV drip and an array of sensor cables attached to it like marionette strings, but only managed to twitch a few fingers. Instead he whispered, "Send her away."

"She claims that a biotechnology start-up she runs in Bangalore has developed medicine that can treat your condition, sir."

A cure? No, it wasn't possible. This woman was playing some angle, telling him what he wanted to hear in order to gain access. Well played. He couldn't afford to refuse her.

"Hello, Uncle Bryce," said the Indian woman in her forties. "It's me, Rohana. You taught me to play chess when I was little, remember?"

Bryce recalled the annoyance of getting stuck watching his business partner's kid while Rajan spent evenings in the lab, so close to their firm's first breakthrough. Back then they couldn't afford a babysitter.

"We were just about to begin clinical trials on this drug when I heard of your diagnosis," she said. "Naturally, we did everything we could to accelerate the process." She held out a small pill. "This isn't a cure, but one of



these per day can alleviate your symptoms and prolong your life by a year or more."

Bryce was sceptical, but he had nothing to lose. With her help, he gulped down the pill.

Every day, Rohana Jethwani would visit and deliver another dose. She never stayed more than a few minutes or said much, but Bryce didn't care because the drug was working. He was getting stronger, feeling better than he had in weeks, beginning to eat solid food. On the seventh day she handed him a sheet of paper along with the pill.

"What's this?" Bryce asked. He was sitting up in bed, reading a quarterly report. He felt strong enough to work again.

"Your bill for the first week."

Aha! Bryce didn't believe in altruism and Rohana's kindness was making him somewhat uncomfortable. He'd gladly pay for treatment. He glanced at the bill and suppressed a chuckle; it was a measly \$127. Like her father, Rohana didn't seem to grasp that pharmaceuticals were always a sellers' market, and consumers would reach as deep as they had to into their pockets when it came to their well-being.

GO NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

"Say, would you consider selling the formula? Or, perhaps, the entire company?"

"I don't think so," said Rohana. "When you taught me chess, you also told me a legend about its creator. Do you remember it?"

Bryce shook his head.

"Some ancient king liked the game so much that he let the creator name his reward. The man wanted wheat: one grain for the first square of the chessboard, then double the amount for each subsequent square. The king agreed, not realizing the enormity of the request."

Rohana stared Bryce in the eye. "You told me that story around the time you 'forgot' to reapply for my father's work visa.

He was forced to move back to India, and to sell you his share of the company mere months before you made millions off his research. He died in obscurity a decade ago, but you didn't even know that, did you?"

Bryce tried to say something, but Rohana cut him off.

"You need one pill per day to live, and I'm willing to supply them. Your first pill was a dollar, the second two dollars, and so forth. It's a pittance now, but your twenty-first pill will cost over a million, and it'll get really expensive after that. In the end, you'll either be dead or I'll own the company you stole from my father. And when I do, you and every other patient will receive care at rates they can afford."

"How dare you blackmail me!" Bryce crumpled the bill in his fist. "I will bring the full resources of Green Industries down on your foolish head."

"This isn't blackmail," said Rohana. "Just a business transaction. Business the way you'd handle it. Going forward, you will wire the money each day and a courier will deliver the pill. Your scientists won't be able to reverse-engineer the formula quickly enough, and if you try anything underhanded, the pills stop for good." She turned to leave. "The next time I see you, I'll either be in charge of Green Industries or attending your funeral. The choice is yours."

She walked away, Bryce still holding the bill in his shaking hand. ■

Alex Shvartsman is a writer and game designer from Brooklyn, New York. Read more of his fiction at www.alexshvartsman.com.

ILLUSTRATION BY JACEY

21 May 2015
Supplement to Nature
Publishing Group journals

21 May 2015

Supplement to Nature

Publishing Group journals

[illegible]

bee care

Busy insects with a big impact

MEET OUR PRIME POLLINATORS

Bees do far more than just make honey. Globally, the 25,000 or so bee species play a crucial part in crop production and in promoting biodiversity. By **Julie Gould**.

TYPES OF BEE

The honeybee (*Apis mellifera*) is the most widely studied bee, yet the approximately 10 species of honeybee comprise less than 0.05% of all known bee species.

Apidae family: 5,811 species

Includes honeybees, bumblebees, stingless bees, orchid bees, carpenter bees and many cuckoo bees.

Halictidae: 4,401 species

Includes many primitively social species as well as sweat bees, which suck perspiration from the skin of humans and animals.

Megachilidae: 4,111 species

Solitary bees, including leaf-cutter and mason bees (the females of which collect pollen with their abdomen instead of their legs).

Andrenidae: 2,952 species

Mining bees that are particularly common in temperate climates.

Colletidae: 2,595 species

Mining and cavity-nesting bees, including plasterer and yellow-faced bees.

Melittidae: 201 species

A small family of mining bees, most of which visit particular flowers.

Stenotritidae: 21 species

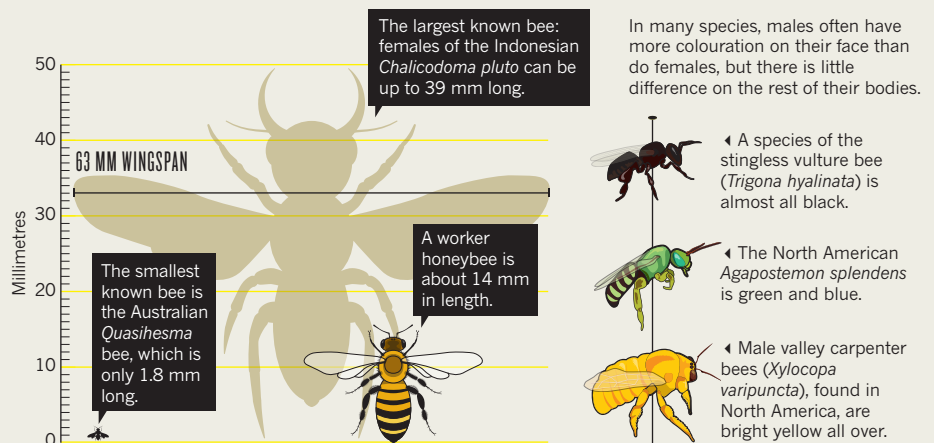
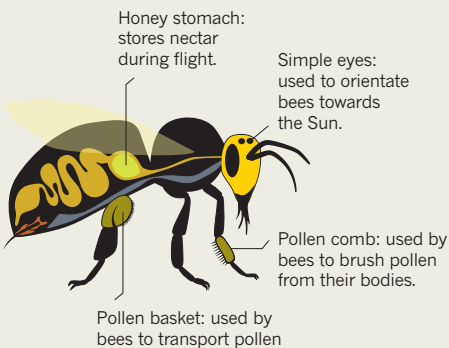
A small family of mining bees endemic to Australia.

5,000

Estimated number of unknown, undescribed bee species. The total number of species is thought to be about 25,000.

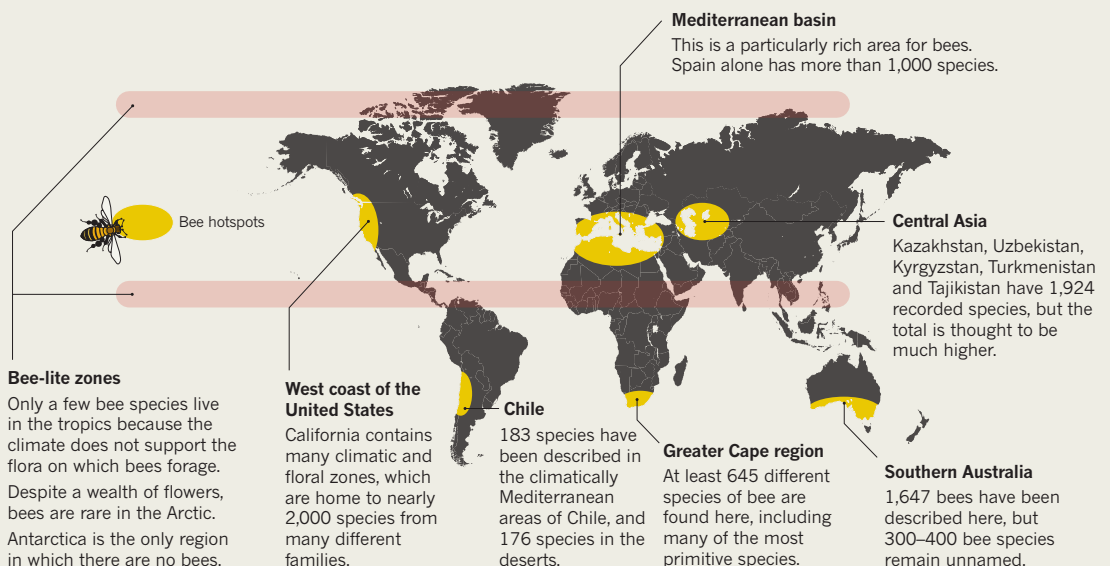
BIOLOGY

Bees, a major group in the order Hymenoptera, evolved from wasps and have adapted to take advantage of the energy available from feeding on pollen and nectar.



DISTRIBUTION

Six global hotspots, all of which have a Mediterranean-style climate, are home to the greatest variety of bee species.



FROM THE BEGINNING ...

For their first few days, all larvae feed on royal jelly (a substance made in glands on the heads of worker bees). Future queens continue to be fed royal jelly; worker bees and drones consume bee bread, which is a mixture of honey, pollen and water.



Eggs hatch after 3 days.

Queens pupate for 12 days; drones and workers for around 7.

In two to three weeks, bees metamorphose.



1 DAY

4 DAYS

8 DAYS

2-3 WEEKS

HATCHING

Young bees emerge every year during the spring.

The number of eggs laid per lifetime varies from eight (or fewer) for some solitary species to more than one million for queens of some social species.

Honeybees are social insects with defined castes: workers (sterile females), drones (fertile males) and queens (fertile females).



DEATH

In a colony, 1-2% of honeybees specialize in removing dead bees, which keeps the colony clean and healthy within the enclosed nests.



8 YEARS

The age of the oldest reported honeybees.



BEE LIFE STORY

The domesticated honeybee is the archetypal bee. Here, we outline the honeybee's life history and highlight some other species of wild bee.

DECLINE

Reliable data are scarce — but they point to a bleak future for many bees worldwide.

300% The global increase in agriculture that depends on animal pollination in the past 50 years¹.

25% The decrease² in domesticated honeybee colonies in central Europe between 1985 and 2005.

60% The decrease in domesticated honeybee colonies between 1947 and 2005 in the United States².

50% The proportion of wild bee species that became extinct over a 120-year period to 2010 in Chicago, Illinois³.

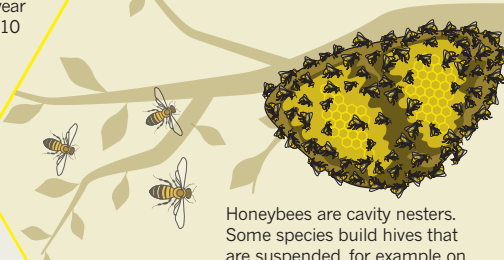
US\$70 billion The estimated global economic value of bee pollination to agriculture per year⁴.

Honeybee workers live on average for three to six weeks between spring and summer. Queen bees can live for up to five years, hibernating through the winters.



NESTING

Hives are not the only homes for bees.



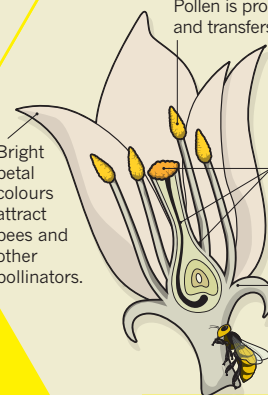
Honeybees are cavity nesters. Some species build hives that are suspended, for example on tree branches or gutters. Other wild honeybee species build nests in hollow spaces such as holes in fallen logs.



Cuckoo bees behave in the same way as cuckoo birds and put their eggs into the nests of other bees.

Pollen is produced on the anther and transfers to bees.

Bright petal colours attract bees and other pollinators.

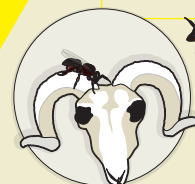


Bees rub pollen from other flowers onto the female part of the flower to facilitate reproduction.

Some plants have oil-producing glands called elaiophores, and specialized bees have evolved to collect oil instead of pollen.

FORAGING

Flowers are the main source of food for bees.

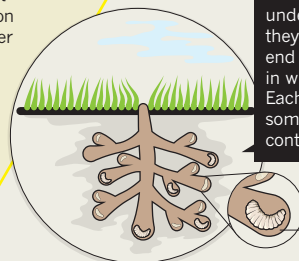


Three species of the stingless *Trigona*, which live in South America, obtain protein from rotting meat.

Most species of bee are vegetarian. Bees rely on pollen for protein, and on nectar for sugar.

Some bees, such as the *Macropis*, collect oils from the plants to line the cells in their nests and make food for the larvae. Adult bees rarely ingest the oils.

Mining bees lay eggs in underground tunnels that they dig themselves. At the end of each tunnel is a cell in which the eggs are laid. Each cell hosts one egg, but some species have nests containing up to 60 cells.



Mining-bee cell



The beeline



100 MILLION YEARS AGO

Bees have been found in fossils (**pictured**) that date to around the same time as the first flowers. “And then there was a huge radiation of plants, accompanied by a radiation of bees,” says entomologist Walter S. Sheppard of Washington State University in Pullman. Today, there are 25,000 species of bees, which are found on every continent except Antarctica (see page S48). Honeybees are members of the genus *Apis*, but they represent only a tiny slice of bee diversity — about ten species. The genus arose around 35 million years ago, probably in southeast Asia. Scientists divide *Apis* into three types: dwarf honeybees; giant honeybees; and, of most interest to us, cavity nesters, which seek out spaces such as hollow tree trunks and fill them with multiple combs. “You can take down the tree they’re in and carry them home,” says Sheppard. And that is exactly what people have done, especially with *Apis mellifera*, which is also called the European, Western or common honeybee. Although its origins are not entirely certain, some scientists believe that *A. mellifera* arose in east Africa — where humans are also thought to have originated.

HUMAN PREHISTORY

Many primates, including gorillas and chimpanzees, are avid honey-eaters. This suggests that honey probably formed part of the diet of our ancestors. “There is a long evolutionary history to the human sweet tooth,” says Alyssa Crittenden, a nutritional anthropologist at the University of Nevada in Las Vegas. Modern hunter-gatherer groups — whose diets are thought to provide clues to the foods that early humans ate — raid wild-bee nests in search of honey. For example, the Hadza people of Tanzania obtain around 15% of their calorie intake from honey. These observations led Crittenden to the provocative hypothesis that this foodstuff may have played a vital part in human evolution, fuelling our energy-hungry, expanding brains. Honey is one of the most calorie-dense foods in nature. From around 2.6 million years ago, early hominins probably had an advantage over other primates in collecting honey — stone blades and axes would have helped them to hack into tree trunks to reach the honey-rich hives of the honeybee.

An Egyptian relief carved nearly 4,500 years ago depicts a bee-keeper working with horizontal hives — these are metre-long tubes made of sun-dried clay that were sealed at one end and stacked like firewood. The relief represents the earliest definitive evidence of bee-keeping, but by the time of the carving beekeeping was already well established in Egypt and possibly elsewhere in the near East, says Gene Kritsky, author of *The Quest for the Perfect Hive: A History of Innovation in Bee Culture* (Oxford Univ. Press, 2010). The hieroglyph for bee (**pictured**) had already been in use for about 500 years.

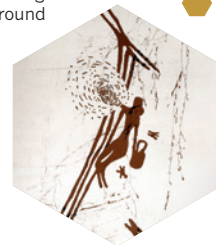


~2450 BCE

BEE ART

~6000 BCE

Evidence of prehistoric honey hunting: a cave painting (**pictured**) made around 8,000 years ago in present-day Spain depicts a person precariously perched on a cliff-face ready to raid a bees’ nest.



LONG-TIME FRIENDS

1000 BCE

Bee-keeping developed independently in several areas of the world.

The Chinese have kept hives of *Apis cerana*, the Asiatic honeybee, for more than 3,000 years. In Central America, the Mayans kept native stingless bees (**pictured**) belonging to the genus *Melipona* — not closely related to *Apis* but also honey-makers — in hollow logs suspended from either the forest canopy or the eaves of houses.



TO BEE...

29 BCE

Roman poet Virgil devotes one book of *The Georgics*, an epic poem on agricultural themes, to bees. He draws parallels between the structure of honeybee colonies and human society.

...AND ALSO TO BEE

1599

In his history play *Henry V*, Shakespeare proves there really is nothing new under the sun by comparing a well-run kingdom to a beehive.

CLOCKWISE FROM TOP LEFT: REDMOND DURRELL/ALAMY; INTERFOTO/ALAMY; VISUALS UNLIMITED, INC./ERIC TOWNREY/GETTY IMAGES; PAUL PEBBING/GETTY IMAGES; NATURAL HISTORY MUSEUM, LONDON

Of all insects, bees — especially honeybees (*Apis mellifera*) — are the most lauded by humans. They have been praised by poets and writers, including Virgil and Shakespeare, and their colonies are seen as a metaphor for human societies. This affinity is no surprise: humans and bees have a long and interwoven history. By Sarah DeWeerd

GRAND VOYAGE

1622

Apis mellifera arrives in the New World. English colonists sent barrels of bees to Virginia on a ship that carried a cornucopia of seeds, fruit trees and other animals.

QUEEN BEE

1942

In the United Kingdom, newlyweds James and Eva Crane receive a bee-hive as a wedding present; the honey was intended to provide a supplement to their Second World War sugar ration. Nuclear physicist Eva Crane (**pictured**) became fascinated with bees and went on to visit more than 60 countries to study the insects, becoming the twentieth century's foremost bee researcher.



HIVE GEOMETRY

1976

Creation of the top-bar hive (**pictured**). This structure was inspired by an ancient Greek design that encourages bees to build trapezoidal combs not attached to the side or bottom of a hive. Practical and inexpensive, the top-bar hive technology has contributed to economic development schemes across the world.



GENOME BUZZ

2006

The first draft of the *Apis mellifera* genome, only the third insect species to have its genome sequenced (Honeybee Genome Sequencing Consortium, *Nature* **443**, 931–949; 2006), provides intriguing hints that, for example, insects tend to have an enormous diversity of smell-receptor genes.

1852

Reverend Lorenzo Lorraine Langstroth, a clergyman from Philadelphia, Pennsylvania, patents the modern box hive, a construction with multiple removable combs that is still in wide use today. The design's interchangeable parts increase the efficiency of managing hives and honey extraction, but may also hasten the spread of disease if contaminated frames are swapped into new hives.

1945 – PRESENT

In the United States at the end of the Second World War, the number of honeybee hives (**pictured**) reaches about five million because honey is used instead of sugar and beeswax is needed for bomb production. Since then, the number of hives has halved, even though the country's agriculture has become more dependent on managed honeybees. Modernized farms needed access to millions of bees to pollinate vast monocultures, but only for short periods of time. These concerns drove the development of migratory bee-keeping. Dennis vanEngelsdorp, an apiarist at the University of Maryland in Beltsville, describes the bees as “a mobile pollination force”. A bee-keeper might transport hundreds of hives thousands of kilometres from Florida citrus groves to New Jersey watermelon fields to Maine's blueberries.



1987

The *Varroa destructor* mite was originally a parasite of *Apis cerana*. It has jumped species and spread to managed *Apis mellifera* colonies around the world. The pest reached mainland United States in 1987, the United Kingdom in 1992 and New Zealand in 2000. “It's like a baby vampire,” Dennis vanEngelsdorp says. The mite transmits viruses and weakens bees' immune systems. The only country to remain free from the *Varroa* mite is Australia.

2006

Bee-keepers notice large numbers of adult worker bees disappearing, emptying healthy hives in a few days. They dub the phenomenon colony collapse disorder (CCD) (see page S52). “We don't have a culprit,” says Dennis vanEngelsdorp, then state apiarist of Pennsylvania who investigated some of the first cases of CCD. Pesticides, stress from moving hives, lack of forage, parasites and disease remain big problems. Although vanEngelsdorp says that he has not seen any confirmed cases of CCD in the past few years, US bee-keepers routinely lose half of their colonies every year.

2015

More revelations from the honeybee genome, first sequenced in 2006, hint at the roots of instinctive behaviours. Innate behaviour, including the honeybee's dance language, is “anything but simple”, says Gene Robinson, a genome scientist at the University of Illinois in Urbana-Champaign. Other work explores the genetics of social behaviour: in May 2015, analysis (K. M. Kapheim *et al*, *Science*, in the press) of the genome sequences of ten species of bees shows that “there are different ways to make a social bee”, Robinson says. Although the genes involved may differ each time sociality evolves, it tends to involve complex gene networks — a pattern also seen in primates.

CLOCKWISE FROM BOTTOM LEFT: BEES FOR DEVELOPMENT; RICHARD JONES/EVA CRANE TRUST; KEN GILLESPIE PHOTOGRAPHY/ALAMY; DESIGN BY MOHAMED ASHOUR/NATURE



THOMAS SPENDER/GETTY IMAGES

A protest against the use of neonicotinoid pesticides takes place outside the Houses of Parliament in London in April 2013.

PESTICIDES

Seeking answers amid a toxic debate

Some see the European Union's ban on neonicotinoid pesticides as a victory for pollinators, but the data suggest that limiting these compounds may do little to stave off honeybee losses.

BY MICHAEL EISENSTEIN

A camera crew followed Dennis vanEngelsdorp's every move as he returned to his laboratory with an unfortunate cargo — packages of dead honeybees. The year was 2006, and the film-makers were chronicling colony collapse disorder (CCD), a newly coined phenomenon in which entire hives die from the catastrophic loss of adult bees. Scientists were still grappling with CCD, but the media had already found a culprit. “As we were opening the first packages, the crew were asking me, ‘This was neonics, wasn't it?’” recalls vanEngelsdorp, then the acting state apiarist in Pennsylvania. “And this was before we had even done anything!”

Neonics — short for neonicotinoids — are insecticides that were introduced in the early 1990s as a more environmentally benign

approach to agricultural pest management. Rather than being sprayed directly onto crops — the common method of applying insecticides — neonicotinoids are typically coated onto seeds, ostensibly limiting opportunities for exposure by bees and other non-target organisms. But traces can be found throughout treated plants, including in the pollen and nectar that bees subsist on. Barely a year after neonicotinoids were first used in agriculture, French bee-keepers were connecting the chemicals with hive losses. The evidence was circumstantial, but a series of apparent bee-poisoning cases in Europe and the United States fuelled the fire.

The resulting debate has seen environmental groups and their supporters rallying around scientists who believe that this class of chemicals contributes to CCD; other scientists think that there is only a minor risk and accuse the press and activists of inflaming fears of ‘killer

nerve agents’. “In Britain, it brought out all kinds of campaign organizations and slogans,” says Francis Ratnieks, a bee behaviourist at the University of Sussex in Brighton, UK, who sees little evidence linking neonics with honeybee losses. In the European Union (EU), the furor culminated in a two-year moratorium on the three most widely used neonics, enacted in December 2013. Still, the extent of harm caused by neonics to bee colonies remains an open question, and two decades of research have yielded as much controversy as clarity.

ALIVE, BUT UNWELL

Ratnieks notes that in past decades, the evidence for pesticide poisoning was unambiguous. “I lived in the US 30 years ago, when there was heavy spraying of insecticides like carbaryl to control insects in sweetcorn,” he says. “We'd see heaps of dead and dying bees

LIFE ON WHITE/ALAMY

SOURCES: LEFT: DOUGLAS, M. R. & TOOKER, J. F. ENVIRON. SCI. TECHNOL. 49, 5088–5097 (2015); RIGHT: THE BEE-INFORMED PARTNERSHIP

in front of hives.” Other, older pesticides were also harmful to humans — most notably, organophosphates, nerve agents that were historically used as chemical weapons as well as to control pests.

By contrast, neonicotinoids have relatively low toxicity in mammals. Furthermore, seed distributors typically apply these treatments before sale, limiting opportunities for overuse. In North America — and until recently, most European nations — neonic seed treatments are widely used on several crops (see ‘No clear pattern’). “In the US, almost no corn is planted without them,” says Christian Krupke, an entomologist at Purdue University in West Lafayette, Indiana. For honeybees, the most important neonic-treated crop is probably oilseed rape (also known as rapeseed or canola), which blooms with bee-attractive yellow flowers. As a source for both vegetable oil and biodiesel, oilseed rape is a highly valuable cash crop in Europe and Canada.

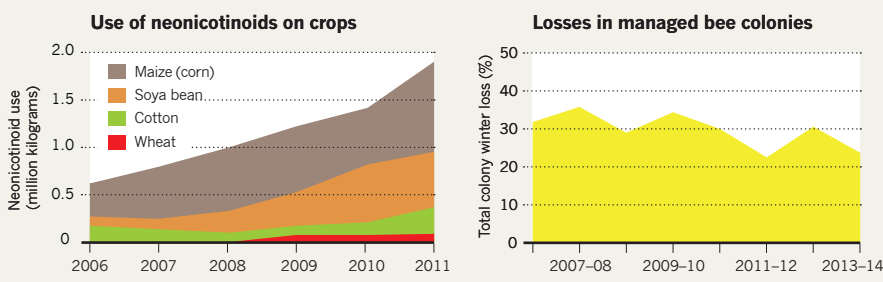
There is no question that all three neonicotinoids banned by the EU — imidacloprid, thiamethoxam and clothianidin — are highly toxic to bees. Treated seeds carry doses that could kill hundreds of thousands of bees. But by the time the crop blooms, only small amounts of the active ingredients are present in the nectar and pollen. Still, laboratory studies have shown that, even at low doses, neonics can have a serious impact on a bee’s brain function. “Neonics affect parts of the brain where sensory information is integrated, including information related to orientation,” says Mickaël Henry, a behavioural ecologist at the French National Institute for Agricultural Research in Avignon. The fear is that these effects essentially confuse bees, making it harder for them to find good sources of nutrition or return safely home with sustenance for their hive-mates.

Henry is among the scientists concerned that prolonged exposure to low doses of neonicotinoids may degrade colony robustness by depleting hives of both bees and food. In a trial conducted in France in 2012, he and his colleagues used radio-frequency identification tags to monitor the homing capabilities of 653 forager bees that were released up to 1 kilometre away from their colony¹. They found that bees treated with sublethal doses of thiamethoxam in sugar water before being released were considerably less likely to return to their hive, with results worse for those foraging in unfamiliar surroundings. Subsequent computer modelling indicated that this steady loss of foragers could jeopardize the hive. “For the first time, we showed that the effects of sublethal doses can lead to indirect mortality, because of bee disorientation, at levels that can put a colony at risk of collapse,” says Henry.

A second study² in 2012, by entomologist Dave Goulson, then at the University of Stirling, UK, reached a similar conclusion regarding the bumblebee species *Bombus terrestris*. Bumblebees differ from honeybees

NO CLEAR PATTERN

The application of neonicotinoids on maize (corn), soya bean and other crops in the United States continues to climb each year. By contrast, nationwide surveys by bee researchers, in collaboration with the Apiary Inspectors of America and the US Department of Agriculture, show highly variable winter losses for honeybee colonies. Equivalent multi-year data are not publicly available for many European countries, although the few data that there are hint at a similar lack of correlation.



in that queens live for only one year rather than several, so annual queen production is crucial to the survival of a colony. For two weeks, Goulson’s team fed bees in 75 colonies with either plain pollen and sugar water or the same foodstuffs containing imidacloprid; for the next six weeks, they observed the colonies foraging freely. The treated colonies showed an 85% reduction in the production of queen bees, and subsequent findings from Goulson³ have implicated impaired foraging. These results highlighted potential real-world consequences for low-level pesticide exposure in bumblebees. “Most studies up to that point had been done with bees in a greenhouse or a cage or even a plastic container, where the bee doesn’t have to be very good at navigating,” says Goulson. Indeed, several other studies have since suggested that non-honeybee pollinator species may be particularly vulnerable to neonicotinoids’ effects (see ‘Plight of the bumblebee’).

“Neonics affect parts of the brain where sensory information is integrated.”

REALITY FIELD

By examining bees in real-world environments, the studies by Henry and Goulson invigorated the neonic debate — indeed, France moved to ban thiamethoxam within months of the publications. But these studies still relied on forced dosing of bees, based on an experimentally determined range of neonicotinoid concentrations — and some experts are wary of their validity. Ratnieks and his University of Sussex colleague Norman Carreck have looked at results from forced-dose trials and found that the studies that showed the greatest risk to bees used doses that were either based on unrealistic or at least worst-case assumptions, and therefore may be of little relevance in field conditions⁴. Ratnieks adds that there could also be considerable variability in the effects of neonics on bees depending on the manner of the dosing. A worker honeybee going about its daily business of picking

up multiple, small doses of neonics in the nectar it collects has a chance to metabolize the insecticide and prevent it from building up as opposed to a bee that has received the same dose applied by a researcher at one time. “Just like if you drink a whole bottle of whiskey in a single session versus having a glass or two per day for a week,” says Ratnieks.

Without clear agreement on how much pesticide bees actually encounter, conclusions drawn from forced-dose studies remain controversial. “There should be increased efforts to do sound studies with real exposure, not just ‘realistic’ laboratory exposure,” says Jens Pistorius, head of bee risk assessment at the Julius Kühn Institute in Berlin, and a bee-keeper himself.

Only a handful of peer-reviewed studies have examined foraging in actual treated crops, and these generally offer little evidence for ill effects in honeybees. One was published⁵ in 2013 by scientists at Syngenta, the Swiss agrochemical company based in Basel that developed thiamethoxam. The study ran for four years in France and examined several indicators of hive robustness for honeybees foraging in either seed-treated fields of maize (corn) or oilseed rape or untreated control fields. “We saw absolutely no effect on the honeybee colonies in those trials, including overwintering success,” says Peter Campbell, Syngenta’s senior environmental specialist, referring to the hive’s capacity to rebound from winter population losses.

A second field study⁶, performed in Canada, reached a similar conclusion in 2014 after comparing bee deaths, honey production and other measures of health in 40 colonies that foraged in fields of untreated or clothianidin-treated oilseed rape. “We are not seeing any impact on honeybees as a result of exposure to canola grown from neonic-treated seeds,” says Cynthia Scott-Dupree, a pest management specialist and toxicologist at the University of Guelph in Ontario, Canada, who co-authored the study.

Both studies have been criticized for conflict of interest. The study that took place in

France was conducted by a neonicotinoid manufacturer, and the Canadian study was financed by one: Bayer CropScience, the German company based in Monheim that developed clothianidin and imidacloprid. “Any research into the safety of agrochemicals should be funded by a government or completely independent entity,” says Goulson, now at the University of Sussex. But Scott-Dupree notes that field trials of pesticides at the scale necessary to yield meaningful results are complicated and expensive, and she bristles at the notion that the funding affected her findings. “It cost us close to a million dollars for a one-year study — where are we going to get that if not from private funds?” she says. “I didn’t get any paybacks to generate data that would support my funders.”

The two studies have also been criticized for inadequate test fields. Both used plots of 2 hectares (0.02 square kilometres) — less than the honeybee’s typical springtime foraging area of 3–12 km², and much smaller than real-world fields. This raises the possibility that bees supplemented their diet with untreated outlying plants, and may have ingested lower doses of the chemical than was assumed.

However, these trials are at least partly supported by field data from Swedish researchers at Lund University — a study funded entirely by government and non-profit foundation resources⁷. The scientists examined the well-being of multiple bee species after the bees had been foraging in untreated or clothianidin-treated fields of oilseed rape with an average size of around 9 hectares. The data showed evidence for adverse effects on the health of bumblebees and other wild bees, whereas honeybee colonies remained largely unscathed. “This doesn’t mean that there aren’t any negative effects on honeybees, but so far I don’t see any evidence from field studies supporting that,” says lead author Maj Rundlöf.

Numerous real-world factors could be mitigating the known toxicity of neonics. Honeybees do not simply binge on their favourite flower. By analysing pollen samples and the waggle dance that honeybees use to communicate the location of nearby food sources to other hive members, Ratnieks learned that bees that live near highly desirable oilseed rape spend barely half their time foraging in the crop⁸. In Britain, he explains, oilseed rape mostly blooms in the spring, when plenty of other flowers offer the bee a variety of food choices. What is more, honeybee colonies can often shake off moderate losses, with enthusiastic springtime reproduction making up for individual deaths — particularly when nutrient-rich crops such as oilseed rape are available. “If you have a steep increase in colony strength, it’s questionable whether you would



Maize (corn) seeds in their natural form (yellow) and treated with neonicotinoids (purple and red).

still find very small effects caused by neonicotinoids,” says Pistorius.

Scientists have also analysed additional data from national bee-health surveys and observations from bee-keepers. Although many keepers continue to report bee die-offs, no clear thread directly links these to neonicotinoids. “In the last decade of having many colonies in treated oilseed rape in real agricultural settings, I have never had a single incident,” says Pistorius. He acknowledges that this observation is anecdotal, but notes that his team’s national bee-monitoring system reports similar findings across Germany, with no incidents apparently associated with neonicotinoid exposure to nectar, pollen or dust drift during sowing from seed-treated oilseed rape since 2005. Carreck, also a lifelong bee-keeper, describes similar observations from the UK government’s Wildlife Incident Investigation Scheme. “There hasn’t been a confirmed incident of honeybees being killed by the approved use of an agricultural pesticide since 2003,” he says.

WHAT BEES SEE

There is one scenario in which the danger of neonicotinoids to bees is unambiguous. In spring 2008, there was an abrupt rise in bee deaths in southern Germany that affected more than 11,000 colonies. The cause turned out to be clothianidin-contaminated dust, abraded from the surface of treated seeds, that became airborne during machine-assisted planting of maize. Similar incidents throughout Europe and North America have also revolved around maize, which is planted in spring when other crops and wildflowers are in bloom — elevating the risk to bees. Krupke, who works in the heart of the US corn belt, began investigating similar reports in 2010. “In all cases, we found that the dead bees had neonicotinoids on them,” he says. His group systematically analysed⁹ the extent of contamination and obtained striking data regarding the seed dust. “It was so incredibly toxic — a bee flying behind a corn planter would just die on the spot,” he says.

After the 2008 incidents, Germany banned neonicotinoid seed treatments for maize. Other European nations required farmers to use deflectors that minimize dust release. In 2013, Bayer CropScience released a new

lubricant for pneumatic planters that limits seed abrasion and hence dust, and the Canadian government mandated its use for treated maize and soya bean. “In last year’s report from Canada’s Pest Management Regulatory Agency, loss of bees due to corn planting had dropped by 70%,” says Scott-Dupree. Unfortunately, neither measure is commonplace in the vast cornfields of the United States; until planting practices change, Krupke is seeking other means to protect bee colonies. “We’re trying to figure out how far from cornfields they have to be before there is no risk of contacting toxic levels of planter dust,” he says.

BANS AND CONSEQUENCES

The central question of whether neonicotinoid seed treatments, when properly applied, are harming honeybees remains murky. Some experts rule out serious danger from the doses found in nectar or pollen: “I just don’t see the exposure being there, and I don’t see the evidence of colony-level effects for honeybees,” says vanEngelsdorp, who is now at the University of Maryland in College Park.

Nevertheless, a cautionary scientific report produced by the European Food Safety Authority (EFSA) in May 2012 — alongside considerable political pressure that included an online petition signed by two and a half million people — moved the European Commission (EC) to action. In spring 2013, following a close vote, the EC enacted a two-year moratorium on imidacloprid, thiamethoxam and clothianidin treatment for bee-attractive crops (including oilseed rape and maize).

Both the initial scientific report and the subsequent EFSA guidance document on assessing pesticide risk have come under fire as politically motivated rush jobs. Unsurprisingly, Syngenta is among the most vocal critics, claiming that the guidance requires pesticide manufacturers to demonstrate safety at a level that is statistically unfeasible. “It requires the ability to detect a 7% effect on honeybee colonies, which is below the natural variability you would see,” says Campbell, adding that the guidance remains the subject of ongoing debate two years after its release. “The EC was using a very controversial, very conservative approach that has not yet been agreed upon within the EU,” he says.

Manufacturers are not the only critics, however; some scientists are calling for a more nuanced approach to pesticide evaluation. “The scientific discussion is often very emotional, and there is also a lot of political pressure,” says Pistorius. “The risk from neonics certainly varies greatly for different routes of exposure, different crops and applications, and this issue requires a substantially more differentiated evaluation that considers these various uses and conditions.”

Conversely, pesticide manufacturers were quick to offer dire predictions of agricultural

THINKSTOCK

disaster that, for now, bear little resemblance to reality. A report funded by Bayer CropScience and Syngenta suggested that, over five years, continued suspension of neonicotinoids could cost the EU between €17 billion (US\$19 billion) and €23 billion. But the data thus far indicate few negative effects — indeed, the EC's crop-monitoring report from December 2014 described a highly productive year for crops such as maize and sunflowers. However, oilseed rape is not included: because of the timing of the planting season, the first harvest of untreated crops will not happen until later this year. Early data from Britain's Home Grown Cereals Authority suggest that only 5% of seedlings in England and Scotland were lost to the flea beetles normally thwarted by neonics. But the report also notes that certain regions were hit especially hard, experiencing losses of 40% or more. "The problem is that it was almost impossible to predict where flea beetle attacks would occur," says Campbell, "and once a farmer knew they had a problem it was too late and the damage was done."

NO QUICK FIX

The EU moratorium may prove to be a missed opportunity for science. Its short duration makes tracking trends a challenge — even if there were a way to detect them. "It's kind of daft," says Goulson. "I don't understand why the EU didn't introduce some measures to at least try to monitor the effects." Furthermore, use of alternative pesticides may be masking any benefits of the moratorium. Several European nations have pursued 'derogations' that allow temporary use of seed treatments — essentially sidestepping the moratorium — and the UK government has authorized the spraying of neonicotinoids. Many farmers are also increasing their use of pyrethroid pesticides, spraying crops up to five times a year rather than just the normal one or two. In addition to also being highly toxic to bees, the route of application for pyrethroids could mean more accidental exposure for non-target organisms such as bees.

Perhaps a greater concern is the illusion that politicians are 'doing something' about bee deaths, but ignoring other important threats. "There's a consensus among bee scientists that long-term declines are primarily due to changes in land use that leave less forage and fewer places to nest," says Carreck. And although vanEngelsdorp's early investigations of CCD-ravaged hives uncovered no straightforward answers, they did reveal high levels of the parasitic mite *Varroa destructor* and widespread viral and fungal infections. Over the ensuing nine years, vanEngelsdorp and many other bee researchers have become increasingly convinced that this mite, which carries diseases and also weakens immunity against other infections, is public enemy number one for bee-keepers. "We've been dealing with *Varroa* mites for a long time," he says. "But

PLIGHT OF THE BUMBLEBEE

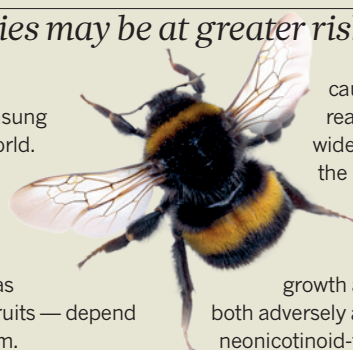
Other bee species may be at greater risk from pesticides

The plump bumblebee, *Bombus terrestris*, is an unsung hero of the agricultural world. Many experts believe that the majority of pollination is conducted by these insects, and that some crops — such as tomatoes and most soft fruits — depend almost exclusively on them.

However, bumblebees may be especially susceptible to the effects of neonicotinoid pesticides. A team led by Dave Goulson, an entomologist at the University of Sussex in Brighton, UK, has found that bumblebees experiencing the neurological effects of these chemicals provide poor support for their hives³. "You get a big drop in the number of bees that come back with pollen, and that's the hive's only source of protein," he says, "so they can't rear enough queens."

Even scientists who are sceptical of the risk to honeybees recognize the importance of assessing potential impact on other pollinators. "Bumblebee colonies are clearly different from honeybee colonies," says Peter Campbell, a senior environmental specialist at Swiss agrochemical firm Syngenta, who is based in Bracknell, UK. "But Syngenta has conducted and submitted for publication a field study that clearly showed no effects of the neonicotinoid thiamethoxam on bumblebees."

Nevertheless, a steady trickle of data over the past few years has given increased



cause for concern. After reanalysing data from a widely criticized 2013 study by the UK Food & Environment

Research Association, Goulson's team found¹¹ that bumblebee-colony

growth and queen production were both adversely affected by exposure to neonicotinoid-treated crops. In parallel, a field study⁷ from the University of Lund in Sweden showed that when foraging in oilseed-rape plants grown from clothianidin-treated seeds, both colony growth and queen production were stunted for the bumblebee, and that solitary bees from the species *Osmia bicornis* failed to build nests. "I don't think any of us expected to see what we saw," says lead author Maj Rundlöf.

Unfortunately, the tests used to assess pesticide toxicity may prove irrelevant to bumblebees — to say nothing of the thousands of wild solitary bee species (see page S62). Given that many of these species are endangered, the problem is all the more pressing. "A honeybee colony has got a huge degree of resilience, but a solitary bee is just that — if a female doesn't lay eggs that hatch, that's the end of her," says Norman Carreck, a bee researcher at the University of Sussex. "We have to look at as many species as we can, because there is quite good evidence that not all of them are as good at detoxifying these substances as others." **M.E.**

today's mite is different from the *Varroa* of 30 years ago, and the viruses have also changed." He adds that neonics are not the only chemicals that bees encounter — he and his colleagues found measurable amounts of more than 120 agricultural chemicals or derivatives in honeybee colonies¹⁰, any of which might affect bee health.

For now, no national ban is under consideration in the United States or Canada, but the issue remains bitterly politicized anywhere there are farms and apiaries. Neonicotinoid manufacturers and advocates claim that environmental organizations are colluding to promote biased research. Environmental (and some bee-keeping) groups accuse researchers who find no clear evidence for neonic harm of being in the pocket of big agribusiness. "We're not funded by these companies," says Ratnieks, "but you have the feeling that if you're not careful, being objective may easily be perceived as being 'pro-pesticide' — which we're not." And in general, the media leave little opportunity for scientists to address these issues with

nuance. Carreck recalls speaking to a British journalist the day the EU ban went into effect. "He wanted a debate, and said he wanted to get 'both sides'. I told him: 'I'm not on anyone's side — I'm trying to be objective here!'" says Carreck. "He immediately lost interest." ■

Michael Eisenstein is a freelance science writer in Philadelphia, Pennsylvania.

1. Henry, M. *et al. Science* **336**, 348–350 (2012).
2. Whitehorn, P. R., O'Connor, S., Wackers, F. L. & Goulson, D. *Science* **336**, 351–352 (2012).
3. Feltham, H., Park, K. & Goulson, D. *Ecotoxicology* **23**, 317–323 (2014).
4. Carreck, N. L. & Ratnieks, F. L. W. *J. Apic. Res.* **53**, 607–614 (2014).
5. Pilling, E., Campbell, P., Coulson, M., Ruddle, N. & Tornies, I. *PLoS ONE* **8**, e77193 (2013).
6. Cutler, C. G., Scott-Dupree, C. D., Sultan, M., McFarlane, A. D. & Brewer, L. *PeerJ* **2**, e652 (2014).
7. Rundlöf, M. *et al. Nature* **521**, 77–80 (2015).
8. Garbuzov, M., Couvillon, M. J., Schürch, R. & Ratnieks, F. L. W. *Agric. Ecosyst. Environ.* **203**, 62–68 (2014).
9. Krupke, C. H., Hunt, G. J., Eitzer, B. D., Andino, G. & Given, K. *PLoS ONE* **7**, e29268 (2012).
10. Mullin, C. A. *et al. PLOS ONE* **5**, e9754 (2010).
11. Goulson, D. *PeerJ* **3**, e854 (2015).

The puzzle in a bee's gut

By analysing bacteria that live in the digestive tracts of bees, researchers hope to learn about the role of microbes in insect health.

BY ALLA KATSNELSON

Sometimes, serendipity arrives on the wings of disease. It was colony collapse disorder (CCD), a mysterious condition that hit honeybee hives in autumn 2006, that brought bees to the laboratory of evolutionary biologist Nancy Moran. Moran, working at the time at Yale University in New Haven, Connecticut, had been studying microbes that live inside aphids and leafhoppers since the early 1990s. Owing to her knowledge of insect-associated bacteria, she was brought in by a team of genome sleuths to analyse RNA samples from sick honeybees. The quest¹ yielded no culprit for CCD, but Moran was surprised to find that whether the bees were healthy or ill, their hindgut (equivalent to the mammalian intestine and rectum) carried a characteristic handful of species that make up 99% of the gut microbial population. “Every single bee had the same bacteria,” says Moran, who is now at the University of Texas at Austin.

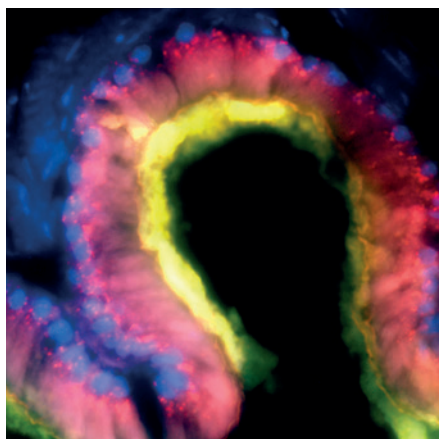
Around the same time, research on the role of microbes in mammalian health was just dawning — helped by a new wave of genomic tools. With her first comprehensive look at the bee microbiome, Moran was hooked. Understanding how these different players help or harm the bee can shed light on bee health, she realized. The small and consistent set of bacteria — unlike the variable species found in the fruit fly *Drosophila*, a common model organism — made studying the microbiome simpler in bees than in mammals. And, because honeybees are social animals that share microbial species, they might provide a good model for studying gut bacteria in mammals too, she says.

COMMUNITY EFFORT

Since Moran's first foray into the honeybee microbiome, a handful of other labs have confirmed her basic finding that all honeybees' guts contain the same core species of bacteria. Bumblebees, although less well studied, also have a characteristic hindgut microbiota, including some of the same species that inhabit honeybees. Honeybees emerge from the pupal stage with a clean hindgut and, over their first three days, acquire the right microbes through interactions within the hive. But researchers have not yet unravelled the mystery of why bee microbiomes are the same across colonies. “There's something special that we still haven't pinpointed in bumblebees and honeybees that allows them to have this very specific microbiota,” says Quinn

McFrederick, an entomologist at the University of California, Riverside.

Scientists are starting to figure out how gut bacteria can affect honeybees, with several microbial roles emerging. The earliest hint of function came from work published in 2011 by Swiss researchers: by raising bumblebees without gut microbiota, they discovered² that the core population of microbes confers protection against trypanosome parasites, which are highly virulent in bumblebees and also present in honeybees. In 2014, Moran's group reported³ that some strains of *Gilliamella* bacteria in honeybees' guts can degrade pectin, a



Cross-section of a honeybee hindgut with staining showing the symbiotic bacterium *Snodgrassella alvi* (yellow), other bacteria (green), insect-cell nuclei (blue) and insect tissue (red).

sugar found in pollen walls. And more recently, researchers in the lab of microbiologist Irene Newton of Indiana University, Bloomington, analysed⁴ RNA from the entire honeybee microbiota to learn more about the microbes' roles. Their key finding was that many species probably help to metabolize carbohydrates — the major constituent of nectar and pollen.

Another lab, led by Jay Evans of the Agricultural Research Service in Beltsville, Maryland, has been exploring the effect of one particular honeybee gut microbe — *Snodgrassella alvi* — on viral infections. At the 2014 meeting of the International Union for the Study of Social Insects, Evans presented research, conducted with Moran's group, showing that the presence of *S. alvi* can cut the number of viruses in a hive by as much as half, apparently by triggering a systemic immune response. Promoting

populations of this species, for example by probiotic supplementation, might improve hive health, Evans says.

Despite the progress in determining function, “there are some very basic things that we don't have answers for”, Newton says — starting with a clearer picture of which organisms are present. Moran's group has shown that there are six to eight core species in the honeybee hindgut. But other species present in fewer numbers could also be playing an important part. Furthermore, ‘species’ is a much more approximate concept in the microbial world than in multicellular organisms (including humans), and genetic diversity among strains within species might be important. To further complicate the picture, Newton's group reported⁵ this year that the queen bee's microbiome is completely different from that of the worker honeybees that have been the focus of most research. This distinction suggests that the queens acquire their microbiota in a different manner from workers.

The gut microbiome is not the only microbial community that is important to bees, says Kirk Anderson, a microbial ecologist at the Carl Hayden Bee Research Center in Tucson, Arizona. Anderson studies bacteria living on the inside walls of the hive, an environment he likens to skin. “These bacteria have evolved to make a living in one of the most extreme antibiotic environments on the planet,” he says — referring to honey and other bee foodstuffs, which are full of antimicrobial chemicals. His work suggests that this external microbiome may have protective effects. “People are looking at gut bacteria,” he says, “but there are all these other populations.”

And beyond even these unknowns, there are many other types of bees to explore. “We've probably looked at some 20 species of bees out of a total of more than 20,000,” says McFrederick. “There's still a lot to be learned.” ■

Alla Katsnelson is a freelance science writer in Northampton, Massachusetts.

1. Cox-Foster, D. L. et al. *Science* **318**, 283–287 (2007).
2. Koch, H. & Schmid-Hempel, P. *Proc. Natl Acad. Sci. USA* **108**, 19288–19292 (2011).
3. Engel, P., Martinson, V. G. & Moran, N. A. *Proc. Natl Acad. Sci. USA* **109**, 11002–11007 (2012).
4. Lee, F. J., Rusch, D. B., Stewart, F. J., Mattila, H. R. & Newton I. L. G. *Environ. Microbiol.* <http://dx.doi.org/10.1111/1462-2920.12526> (2014).
5. Tarpy, D. R., Mattila, H. R. & Newton I. L. G. *Appl. Environ. Microbiol.* <http://dx.doi.org/10.1128/AEM.00307-15> (2015).



ENTOMOLOGY

The bee-all and end-all

Seven scientists give their opinions on the biggest challenges faced by bees and bee researchers.

ROBERT PAXTON

Honeybee viruses

Head of general zoology, Martin Luther University Halle-Wittenberg, Germany

Honeybees are declining in number across the Northern Hemisphere. There is broad consensus within the scientific community that their most serious threats are pathogenic microbes, particularly viruses, and the parasitic mite *Varroa destructor*, which transmits viruses while sucking the blood of the bee. A major challenge is to show whether *Varroa* mites also lower the immune response of the host bee to these viruses. Or do the mites provide an environment that selects for better-replicating or more-virulent viral

variants? — or both.

Honeybees host more than 50 types of microbe, which next-generation sequencing technologies are helping us to explore. Researchers are trying to determine which microbes are pathogens and how to control them. We need to understand how pathogens interact with other stressors — pesticides and poor nutrition — in ways that harm honeybee populations. The field would benefit from mechanistic models describing these interactions at the molecular level, revealing targets of selection for host tolerance or pathogen suppression.

The impact of pathogens on individuals may not translate into colony-level effects. So it is crucial that we figure out where the line is — how many individual losses and in what season — will destroy a colony. Acquiring this understanding will involve both empirical studies and theoretical modelling.

Perhaps most importantly, governments

need to regulate the movement of disease-carrying honeybees to reduce the invasion and emergence of new pathogens. The same honeybees that are imported to support pollination and agricultural production also threaten native pollinators (as well as other honeybees) and hence undermine sustainable provision of these ecosystem services.

MARK BROWN

Diseases in wild bees

Professor of evolutionary ecology and conservation, Royal Holloway, University of London

The 25,000-or-so species of bee are important components of biodiversity and are essential for pollinating crops and wild plants. Although we have limited data, it seems that the populations of many of these species are in decline. Throughout the twentieth century, the major driver of the decline in the number of bees was habitat loss, but since then the threat posed by new diseases has come to the fore.

New or emerging diseases are linked to the rapid declines over the past 20 years in North American bumblebee species and to the dramatically shrinking range over the past 5–10 years of a charismatic South American bumblebee, *Bombus dahlbomii*. Commercial bumblebees that are bred and used for pollination have been blamed as the source of the diseases. Wild bumblebees are also susceptible to an array of viruses that are common in managed honeybees, with some viruses showing patterns suggestive of spread from managed honeybees to wild bumblebees¹.

Although there is evidence that parasites can be transmitted from commercial and managed bees to wild bees, we lack proof that these parasites cause a decline in the number of wild bees. Specifically, we have not definitively identified the direction of transmission for parasites and pathogens, and we have little idea of the impact they have on wild-bee populations.

In the meantime, given the importance of wild bees, application of the precautionary principle is justified. Researchers should support commercial producers of bumblebees in generating disease-free colonies, and governments should ensure that the use of commercial bees is limited to escape-proof greenhouses, such as those that are used in Japan. In addition, the export of commercial bumblebees to countries where the commercial species is non-native should be banned. Similarly, disease management in honeybees needs to be supported, to protect both the

honeybees and the wild bees with which we know they share diseases.

Emerging diseases are a global problem for biodiversity. We need to grapple with them in our wild bees to reverse current declines and to prevent future disasters.

MICHAEL KUHLMANN

Expertise in decline

Head of insects division, Natural History Museum, London

When it comes to bees, Europe is the best-studied continent and has nearly 2,000 known species. Even so, more than 120 new species of European bee have been described since 1990, and there are probably another 100–200 species still to go. As we build up this knowledge base, we need to do more than just create an inventory: we need to explore the diversity of bees' life histories and flower specializations to develop effective conservation measures and assess the impact of climate change.

Bee taxonomy is notoriously difficult. Many species — often common ones — look very similar to each other yet have different life histories. Taking on this daunting assignment is a small and ageing cadre of skilled taxonomists. Shortage of taxonomic expertise has already left some European bee genera orphaned — without any specialists working on them — and is a serious bottleneck for the rising demand of bee identification in pollinator research. This taxonomic crisis was exposed in the assessments prepared by the International Union for Conservation of Nature and published in 2014 in the *European Red List of Bees*². For 80% of the bees on the list, population trends are unknown, and more than half of all species were labelled as 'data deficient', making it impossible for even an indirect assessment of risk of extinction.

Technology, including DNA barcoding,

might help. It can accelerate the speed of identification of recently collected bees, but is of limited use for old museum collections in which the DNA has degraded. Furthermore, in the parts of the world where this technology would be of most use, such as Asia and Africa, the lack of even basic taxonomic information often makes identification impossible.

Taxonomy urgently needs investment and, crucially, training and encouragement of young academics if we are not to lose an invaluable treasure of expertise.

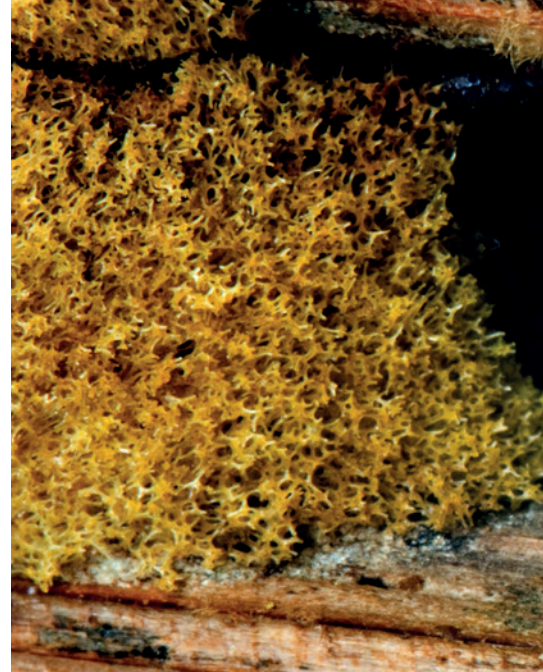
DAVE GOULSON

De-intensify agriculture

Professor of biology, University of Sussex, UK

Bees are often described as the 'canaries in the coal mine' when it comes to the health of the environment. Intensively farmed land is a hostile environment for bees: there are few flowers or quiet places to nest, and many pesticides. We tend to accept that such practices are necessary to feed the growing human population, but we should challenge that assumption.

An ideal farming system would sustainably produce sufficient amounts of healthy food yet also minimize adverse environmental impact. Modern intensive farming fails abysmally to satisfy these basic criteria. For example, around the world about 100 billion tonnes of soil are either degraded or washed away each year, which is clearly not sustainable³. Modern farming is highly dependent on artificial fertilizers, which contribute substantially to climate change. Biodiversity is declining at an unprecedented rate. The loss of bees has attracted attention because our food supply directly depends on these insects, but the reduction in their population is symptomatic of a much broader problem. Most wildlife associated with farmland is



also in decline, including birds, butterflies and beetles.

The majority of investment in agronomic research comes from industry, and tends to focus on increasing yields. Yet we already grow enough food to feed the projected global population of 9 billion in 2050 — we just waste an awful lot of it. We need investment in research and support for sustainable farming systems with reduced inputs — systems that conserve the soil and minimize the impact on wildlife such as bees. Industry is unlikely to invest in ways to reduce inputs, for supply of those inputs provides much of its profit. Surely it is the role of government to intervene in this situation. Do we really want to trust big business to shape the future of farming, and to look after our bees?

AXEL DECOURTYE

Listen to the beekeepers

Scientific director, French Technical Institute of Beekeeping and Pollination, Avignon, France

In response to the inexplicable losses of honeybee colonies in the past two decades in Europe and the United States, research has been focused on understanding the underlying causes. Papers that have been published during this time account for nearly 45% of all publications on the honeybee. Although it is appropriate to try to understand how to act, time is running out.

The main causes of honeybee colony loss have been identified: the parasite *Varroa* and associated viruses; pesticides; and food shortage in the form of wildflower loss.



Many bee species look identical, yet the number of taxonomists who can tell them apart is in decline.



A honeybee leaves the hive unwittingly carrying a *Varroa destructor* mite (inset).

These factors provoke a complex cascade of events, often with delayed effects on bee population dynamics⁴. We cannot afford to continue focusing only on identifying the precise mechanism that is driving the losses while ignoring beekeepers' calls to invest in addressing the known causes.

We need an approach that better integrates the needs of beekeepers, farmers and scientists — for example, by establishing teams drawing on all of these communities to undertake the research and development. Top priorities include the breeding of *Varroa*-resistant bees. The efficiency of the breeding selection also depends on the creation of quality-control procedures by breeders of queen bees. This point is critical to beekeeping sustainability — a new but growing concept. Government policy initiatives should support the remodelling of farms to improve food sources for bees in the landscape and to reduce the use of pesticides. Although the neonicotinoid moratorium in the European Union was a welcome move in this direction, it is not sufficient and addresses only one of the known causes.

JEAN-MARC BONMATIN

Stop poisoning farmlands

Research chemist, French National Centre for Scientific Research (CNRS), Orléans, France

Wild and managed bees are facing an unprecedented situation in which their environment and food resources (pollen, nectar and water) are becoming contaminated by cocktails of pesticides at levels known to have adverse effects. We need to

find ways to reduce bees' exposure to these pesticides, which are mainly insecticides and fungicides.

Of particular concern are neonicotinoids, known as neonics. Laboratory studies have shown that these systemic neurotoxins directly affect bee health and colony performance. And combinations of neonics with other insecticides and fungicides, as well as with certain infectious agents, act together in the bee to amplify the negative effects.

Neonics represent one-third of the global insecticide market; they are used by growers of grains, vegetables and fruit, as well as to kill livestock parasites such as lice and fleas.

"The prophylactic and extensive use of neonics is the major cause of the decline in bee populations."

The prophylactic and extensive use of neonics, combined with their very high toxicity to invertebrates, persistence in soils and solubility in water, is the major anthropogenic cause of the decline in bee populations over the last two decades⁵. And bees are not the only victims of neonics: these pesticides are also harmful to terrestrial and aquatic invertebrates, birds and fishes, both directly and through the food chain⁶.

Although three neonic insecticides have been restricted in Europe since 2013, certain prophylactic uses are still allowed — or other insecticides are applied in their place. The use of insecticides as an insurance policy conflicts with the European Commission-mandated policy of integrated pest management; a directive issued in 2009 states that pesticides should not be used for prevention, only as a last resort. The burden on pollinators will decrease only when pesticide use does, and this will only occur when integrated pest management becomes standard practice in farmlands.

PAT WILLMER

Too many commercial hives

Professor of biology, University of St Andrews, UK

One major concern for bee conservation is the stress introduced into the whole pollination system by having too many commercial honeybee hives. The intensive management of *Apis* hives by industrial beekeepers magnifies all other problems.

Regular long-range transportation of hives to service seasonal orchards can stress and disorientate their inhabitants. Honeybees' health may also suffer from the low pollen diversity found in crop monocultures. Out of season, or while in transit, the bees' nutritional needs are poorly served by maize (corn) or grape syrup, or fructose solutions, which are substituted for the richer honey that has been harvested. Finally, honeybees' natural reproduction is limited because new commercial hives are typically started using artificially inseminated queens, a practice that reduces genetic diversity.

Together, these issues lead to increased pest and pathogen problems; agrochemicals applied to attempt to control them can worsen matters in the longer term, because miticides and antibiotics (plus herbicides and insecticides that bees bring in from foraging) may affect bees' gut microbiota, and reduce the insects' ability to adapt to their pests. All these issues are amplified in *Apis* because the bee's genome has evolved to contain few detoxification and immunity genes, presumably reflecting the low toxin content of nectar and pollen, and the social behaviours that confer some antimicrobial protection.

A reduction in the commercial honeybee population, whether deliberate or from colony collapse disorder, may not be a bad thing. We should use the opportunity presented by the fall in commercial beehives to support native wild bees and encourage natural honeybee-keeping, while providing enough floral diversity so that the bees we do have can collectively provide full and balanced pollination services. ■

1. Fürst, M., McMahon, D. P., Osborne, J. L., Paxton, R. J. & Brown, M. J. F. *Nature* **506**, 364–366 (2014).
2. Nieto, A. et al. *European Red List of Bees*. (European Union, 2014). Available at go.nature.com/c4g8lm
3. Govers, G., Van Oost, K. & Wang, Z. *Procedia Earth & Planet. Sci.* **10**, 313–318 (2014).
4. Becher, M. A., Osborne, J. L., Thorbek, P., Kennedy, P. J. & Grimm, V. J. *Appl. Ecol.* **50**, 868–880 (2013).
5. Sánchez-Bayo, F. *Science* **346**, 806–807 (2014).
6. Bijleveld van Lexmond, M., Bonmatin, J.-M., Goulson, D. & Noome, D. A. (eds) *Environ. Sci. Pollut. Res. Int.* **22**, 1–154 (2015).



Honeybees use the 'waggle dance' to convey information about the location of a food source to a hive.

ANIMAL BEHAVIOUR

Nested instincts

The many levels of bee behaviour offer insights on everything from population dynamics to molecular changes.

BY LAUREN GRAVITZ

Western honeybees (*Apis mellifera*) live in highly complex societies, running nurseries and coordinating food searches that can take them kilometres away from their hives. And they do it all without leadership. "It's like the lights are on but nobody's home," says Gene Robinson, an entomologist studying bee genomics at the University of Illinois Urbana-Champaign. Despite the presence of a queen bee, the hive is not quite so autocratic as her title suggests: a queen's primary function is to lay eggs, which

develop into the male drones and female worker bees that populate the colony.

Honeybees' leaderless organization provides researchers with insight into unrelated systems that have similarly decentralized control — such as the brain or the stock market. "They are compelling models for the study of social life and social behaviour," Robinson says. "They live in highly complex societies that show extreme forms of integration, cooperation and communication."

A huge part of the honeybee's appeal is that researchers can study the insect's behaviour at every level of biology, from their overall social

structure to the minutia of genetics and epigenetics — as well as the interplay between these levels. Only with this complete picture can researchers hope to understand how such simple insects coordinate their behaviour so precisely and with such complexity. "The bee is well suited to address these questions at different levels of biological organization," Robinson says. Every aspect of the hive can be studied to better understand different biological systems, from how genes and epigenetics affect a single bee, to how individual behaviour can affect dynamics of the entire hive. "We have a nested Russian-doll model," says Robinson. Each doll can be studied as a separate entity yet fits neatly into the whole. And with this model, researchers hope to tease out the many layers of a complex system.

HIVE MIND

The outermost doll — the one that faces the world — is the hive itself. Bees manage their colonies by means of specific divisions of labour in which worker bees specialize in different roles. After emerging from their pupae, worker bees typically spend their first few weeks inside the hive, caring for the larvae and performing other housekeeping tasks. After that, they graduate to become foragers, and spend the next few weeks searching for pollen and nectar then sharing the locations of these treasures with the rest of the hive. The transition from hive bee to forager is not dictated by age alone — a hive has to maintain balance among its types of workers, and so bees can speed up, slow down, or even reverse this process as necessary. "If you have old forager bees in a colony, their presence inhibits young bees from becoming foragers," says Andrew Barron, who studies the neurobiology of bee behaviour at Macquarie University in Sydney, Australia. "Without them, younger bees transition faster."

Immature bees are less-effective foragers than are older bees. Barron's work has shown¹ that bees that began foraging before they were two weeks old spent less time outside the hive and went on fewer foraging flights, yet spent more time on each flight. And the younger the bee, the less likely it was to survive beyond 30 minutes outside the hive.

Thus, the hive maintains a balanced ratio of forager bees and hive bees. But if stressors (such as disease or pesticides) kill foragers at too high a rate, younger and younger bees enter the foraging force. When too many bees begin to forage prematurely, the amount of food brought back to the hive declines. And, as fewer foragers survive, new workers mature even earlier, creating a vicious cycle that can cause the colony to collapse. "The colony has a tipping point," Barron says. Quickly replacing foragers with young bees allows a hive to buffer stressors up to a point, he explains. "When that buffer is exhausted, the colony is in dramatic trouble."

The actions of individual bees can substantially alter the health of the hive. Because a hive

HEATHER BELL

usually has several available food sources, a colony has to allocate foragers appropriately. One of the honeybee's most distinctive behaviours is the waggle dance: a series of movements, performed by a forager on her return to the hive, that convey information to all nearby bees about the direction, distance and quality of a food source.

LANGUAGE OF THE DANCE

Because each food source varies in quality over time, bees also need a way to communicate when it is time to stop visiting a site. To convey this message, bees have developed another dance move that serves as a stop signal: a bee butts her head up against a waggle dancer and vibrates at just the right frequency to halt the waggler in her tracks. James Nieh, who studies bee behaviour at the University of California, San Diego, published his findings² on the stop signal in 1993. He has since determined that the bees giving these stop signals are foragers who have already been to the place the waggle dancers are promoting and discovered it to be less than ideal: perhaps the food is gone, the site is overcrowded or the bees were attacked.

In a poster at the Animal Behavior Society meeting in 2011 in Bloomington, Indiana, Nieh described how attacks by spiders, wasps and even a lab-made 'robo-predator' (forceps attached to a spring and activated by a switch), all elicit stop signals from forager bees. The same motion can also help the colony in choosing a new nest site — the fewer stop signals given to dancers describing a site, the more appealing the location. "It's a complex signal used in complex ways in very different circumstances," Nieh says. "Yet the effect of the signal in all these instances is the same: it causes all these waggle dances to stop."

Echoing Robinson's nested-Russian-doll comparison, Nieh describes the signals between individual bees as being analogous to the communication between brain cells. Just as neurons can excite other neurons, the waggle dance acts as an excitatory signal to stimulate action and foraging. And the stop signal acts much like one neuron inhibiting the signal of another. Which in effect, he says, "is how the entire brain or the entire colony achieves a complex decision".

MARKING LINES

It is in the DNA where the most fascinating — and the tiniest — of the Russian dolls sits. This is a realm that researchers are probing even deeper. Benjamin Oldroyd, who studies the behavioural genetics and evolution of honeybees at the University of Sydney, describes how bees engage in a battle of the sexes. Each queen bee mates with upwards of 20 drones, usually within a period of just a few days, and then she stores the sperm for use throughout her lifetime.

Anything that gives one male's daughters a greater chance of becoming a queen or a

reproductive worker enhances his genetic legacy. But to ensure reproductive harmony among her worker daughters, a queen should ideally be able to thwart any such attempts at manipulation by the drones. Given these opposing strategies, males and females within any given bee subspecies typically evolve together to reach an equilibrium in which neither sex has a distinct genetic advantage.

Cross-breeding two bee subspecies can help to show the extent of the evolutionary changes in sexual one-upmanship. In a study³ published in 2013, Oldroyd crossed two subspecies of African honeybee: the Cape honeybee (*Apis mellifera capensis*) and the African (or killer) honeybee (*Apis mellifera scutellata*). The

"If you have old forager bees in a colony, their presence inhibits young bees from becoming foragers."

female *A. m. capensis* typically has more ovarioles (components of an insect's ovaries) than do female *A. m. scutellata* bees. In theory, when the two subspecies are crossed reciprocally, it should

not matter which subspecies is the father: offspring from both crosses should have the same number of ovarioles and hence the same reproductive capacity. But that was not what he found. Rather, one group of daughters had a distinct advantage: those with an *A. m. capensis* father had about one-third more ovarioles than did those with an *A. m. scutellata* father, indicating that the *A. m. capensis* males were employing a trick that extends beyond pure genetics to improve their daughters' fertility.

Oldroyd thinks he knows how the *A. m. capensis* bees are gaining the advantage. "It suggests that males are putting epigenetic marks in their sperm to try and increase the genetic success of their offspring," he says. Epigenetic alterations are chemical tags such

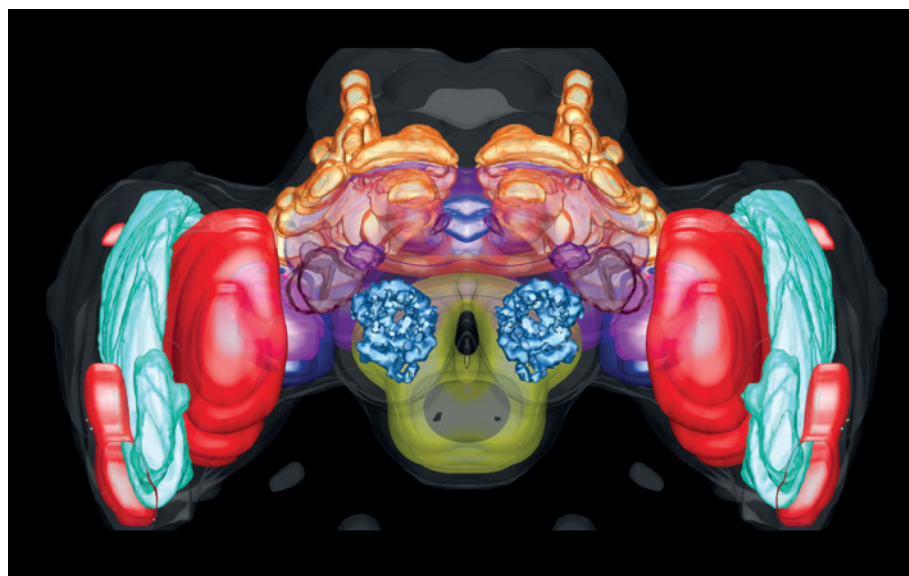
as methyl groups that are added to or removed from genes to turn them off or on. The honeybee was the first insect found to have a fully vertebrate-like methylation system, a fact that has led many researchers to hunt for ways in which bees may be using epigenetics to help further the success of their own lineage. If, as Oldroyd suspects, *A. m. capensis* males are epigenetically tagging the genes in their sperm in a way that enhances their daughters' fertility, then these are tags that the *A. m. capensis* females are able to counteract, but the *A. m. scutellata* females cannot.

Oldroyd is looking for additional evidence that this process involves an epigenetic mechanism. In a more recent experiment, not yet published, he removed an *A. m. capensis* queen from her hive, which led female *A. m. capensis* worker bees to start laying eggs. He found that these single-parent worker eggs have more methylation in their genome than *A. m. capensis* dual-parent eggs. This indicates that there are epigenetic mechanisms in play. Oldroyd says that he still needs to find the specific genes that are being methylated and tease out the effects, but is hopeful that he will be able to find proof of the first epigenetic battle of the sexes in insects.

Examining layer after layer of the honeybee Russian doll gives researchers more than just insight into bee behaviour. Within the hive, says Robinson, are "all the traits that are important to us in understanding complex systems, whether our own society or our own bodies". ■

Lauren Gravitz is a freelance science journalist based in Hershey, Pennsylvania.

1. Perry, C. J., Søvik, E., Myerscough, M. R. & Barron A. B. *Proc. Natl Acad. Sci. USA* **112**, 3427–3432 (2015).
2. Nieh, J. C. *Behav. Ecol. Sociol.* **33**, 51–56 (1993).
3. Oldroyd, B. P. et al. *Proc. R. Soc. B* **281**, 20132388 (2013).



Just like the hive, the honeybee brain has decentralized control.



A male mason bee, *Rhodanthidium sticticum*, makes its nest in the empty shell of an *Otala* snail.

WILD BEES

Lone rangers

Solitary bees receive scant attention, but research shows that they are vital pollinators of crops and wild habitats.

BY LUCAS LAURSEN

In a green field outside Madrid, at the foot of the snow-covered Guadarrama mountain range, lies a sun-faded snail shell. Its opening sealed with a cap of dried mud, the shell contains the larva of a wild, solitary bee, together with its first meal of bee bread — a mixture of pollen and nectar. Entomology graduate student Daniel Romero picks up the shell and, concluding that it contains the nest of a mason bee, stores it in a clear plastic tube, labels the red cap with a marker, and closes it.

Back at the Complutense University of Madrid, Romero sets ten tubes of the nesting bees he collected on his professor's desk. They are just a fraction of the hundreds of samples that he and his colleagues will gather during a four-year Spanish government-funded study of how artificial chemicals are affecting the biodiversity of wild pollinators and their immune and reproductive systems. In the warmth of the office, some of the young adults twitch and scratch at their now-crumblily mud

doors. Researchers watch the young adult bees slowly emerge into their new world. When the air cools and the humans leave the room, the bees return to their pollen pillows. Unlike honeybees, solitary bees buzz to their own drum.

Wild bees perform vital pollination work, but their lives are mostly a mystery. Romero and a handful of other researchers around the world are helping to fill in the blanks. And — given the threats posed to all bees from pesticides, parasites, pathogens, climate change and habitat loss — a growing number of scientists, farmers and regulators are asking how we can better protect wild bees and improve their contribution to agriculture and ecology.

A SOLITARY SPECIALIZATION

Wild bees constitute the majority of the estimated 25,000 species of bees (see page S48). Some are social insects, including feral honeybees and the familiar bumblebees, but most are not.

Unlike honeybees (*Apis mellifera*), which are

indiscriminate pollinators, solitary bees have co-evolved relationships with specific flowering plants: some feed on only one species, and have probosces and leg- and body-hairs that have adapted to the shape of their favourite flowers. As a result, solitary bees can be more efficient pollinators of these favoured plants than honeybees. The diversity of solitary bees also means that they have a different range of tolerances for environmental conditions, such as temperature, wind speed and the number of daylight hours — so a farmer or conservationist managing the needs of these bees requires a larger range of techniques.

The life cycles of solitary species share some similarities: the insects collect and eat pollen, some of which they store for their offspring to eat when they hatch. But most solitary bees do not produce honey, which means that — until the late twentieth century — humans intent on agriculture more or less neglected them.

Hence Romero's fieldwork. He and his supervisor, entomologist Concepción Ornos, spurred by worries about the decline in honeybee and bumblebee populations, started a project to examine how susceptible solitary bees are to the same hazards. Some of the work is fairly basic. "In Spain, we're still learning what bees there are," Ornos says. The Iberian peninsula hosts nearly 1,000 of Europe's estimated 2,000 bee species, of which only a handful are well-described in taxonomic literature. The European Red List of Bees (A. Nieto *et al.* *European Red List of Bees* (European Commission, 2014); available at go.nature.com/c4g8lm), showed that both the diversity of bee species and the number with incomplete descriptions is higher in Spain than anywhere else in Europe. After identifying species, Ornos's team must map their distribution before they can examine how healthy the populations are and what environmental risks might threaten them. Then the real work begins: trying to quantify the impact of different kinds of environmental pollutants on wild-bee health and biodiversity.

GOVERNMENT INTEREST

Efforts by researchers to learn more about wild bees are starting to inform policy. Three years ago, the European Food Safety Authority (EFSA) commissioned a scientific working group to help draft risk-assessment guidance that, for the first time, considers honeybees, bumblebees and solitary bees separately. The updated guidance, published in 2013 but still under review, could arm regulators with the information they need to include wild bees (both social and solitary) in their land-management plans and tests for new pesticides. Right now, the rules treat all bees the same, yet practices designed to protect honeybees may not prevent wild bees from absorbing toxic chemicals, given their very different life cycles. "You can't really extrapolate from *A. mellifera*," says entomologist Fabio

DANIEL ROMERO

T. L. PITTS-SINGER/USDA ARS POLLINATING INSECTS RES. UNIT, LOGAN, UTAH

Sgolastra of the University of Bologna in Italy, who co-authored the scientific opinion that underpins the EFSA guidance document.

Sgolastra expects objections from farmers and pesticide companies because the new risk-assessment guidance requires more tests to ensure the lowest risk possible to the health of all bee species. “It’s quite conservative,” he says. Still, Sgolastra argues that although these tests might be expensive, in the long run they will pay off. A UK study found that honeybee pollination capacity is falling: the insects supply only one-third of demand, the remainder being picked up by wild pollinators including bees. And in the United States, wild bees were the most frequent visitors to three out of four crops studied. Growers and governments are now struggling to strike a balance between protecting wild pollinators and maintaining existing production levels.

The United States, too, is looking to protect its pollinators. In June 2014, President Barack Obama initiated a pan-agency review, which sent researchers across the country scrambling for evidence of the health or otherwise of its pollinators, says US Department of Agriculture (USDA) entomologist Theresa Pitts-Singer. The draft report, expected to be published this spring, will probably call for more basic research on wild-bee populations and how pesticides affect them. Whether these recommendations will be implemented will depend on how much money Congress allocates in response — the task-force report has no direct regulatory power. Pitts-Singer — who works in Logan, Utah, at the only USDA laboratory dedicated to studying the use of solitary bees in agriculture — is hopeful that this could bring the needed funding for solitary-bee research in the United States. “If we’re going to get it,” she says, “it’s going to happen now.”

AGRICULTURAL APPLICATIONS

Agronomic research on wild bees has a short history. After the Second World War, Japanese farmers began managing wild-bee populations by building portable nests, replacing honeybee populations devastated by overuse of pesticides such as DDT. US researchers began promoting the practice to US farmers in the 1970s, using both domestic and Japanese species of wild bees.

The USDA provides some research support for farmers for the alfalfa leafcutter bee, *Megachile rotundata*, and for the blue orchard bee, *Osmia lignaria*, both of which are solitary. But these two wild-bee species are the exceptions. Entomologist Jordi Bosch of the University of Barcelona in Spain has worked in the



A purpose-built shelter with polystyrene nesting boards for the alfalfa leafcutter bee, *Megachile rotundata* (inset), in an alfalfa field in Utah.

United States to develop techniques for managing these and other species of wild bee. He says many research questions remain about which types of wild bees interact with which species of agricultural flowering plants — “And above all, why?” he asks.

Researchers are also looking at how to make portable nests that can guarantee sufficient pollination productivity to be worth nurturing wild bees and even transporting them from one pollination site to another. Researchers at the USDA’s bee laboratory in Logan have tried varying the temperature during larval development to see what effect it has on the timing of the bees’ emergence from their nests, and on their survival rates. And there is a small but growing number of private operations that offer solitary bees for sale. But these businesses are nothing like as reliable as honeybee production: the number of bees available from season to season is unpredictable and the pollination performance of any given batch depends heavily on environmental conditions and release techniques.

KEEP IT WILD

Although some species of solitary bees are suitable for management, researchers shy away from the word ‘domestication’. Many wild bees have value in the untamed landscapes where they pollinate wild plants and help to anchor ecosystems in ways that provide indirect benefits, such as habitat for other insect species and suppression of pests.

Wild-bee populations are exposed to the same dangers as honeybee colonies, but with less protection. As industrial-scale farms plant larger and ever-more orderly swathes of the same few crops, it becomes harder for wild bees to find suitable flowers to forage in or to make nests in untilled ground or in, stray detritus. Monocropping is simplifying the diets of

wild bees. If these bees had a more diverse, complex diet, asks Pitts-Singer, would that help to buffer them against pests and disease?

When researchers have a better understanding of how modern agriculture affects wild bees, they can start to explore improvements in landscapes and practices. The USDA has funded a four-year study to examine the best way to optimize farming landscapes to promote native and managed pollinators. The answers may be surprising: a 2015 study across 36 sites in the United Kingdom recorded a higher variety of bee species in urban areas than in farmland (K. C. R. Baldock *et al. Proc. R. Soc. B* **282**, 20142849; 2015). Working out the optimum ecological balance for wild bees involves collecting them in the real world, as Romero is doing in Spain. It also requires expensive chemical analyses of wild bees’ food sources, nests, pollen and bodies to follow the flow of food and agricultural chemicals through these increasingly artificial ecosystems. “To track it all the way through, it’s just insanely hard and expensive,” says Pitts-Singer: each sample can cost US\$150–200 to analyse with a mass spectrometer. And that is assuming you can even find the insects: “You have no idea what’s happening in the wild lands,” says Pitts-Singer. Wild-bee nests are hidden in the ground, tucked into tree stumps, burrowed into beetle tunnels, and inside snail shells.

Yet it is their wildness and diversity that makes solitary bees so valuable. They live alongside our existing agricultural systems, with different vulnerabilities and strengths from domesticated honeybees. And they promise to inject our agricultural landscape with a healthy dose of biodiversity — and pollen. ■

Lucas Laursen is a freelance journalist based in Madrid.



PAUL D. SAMUEL, DAE DALUS FLIGHT SYSTEMS, LLC & J. SEAN HUMBERT, UNIV. MARYLAND

An imitation insect, called a 'flapping wing micro air vehicle' does not resemble a bee, but it can hover in the air, turn rapidly and take off easily.

AERODYNAMICS

Vortices and robobees

A growing understanding of insect flight is helping scientists to build tiny flying robots.

BY NEIL SAVAGE

The flight of the bumblebee is a remarkable feat. These bees and related species can fly long distances to find flowers, then pause and hover in place, shrugging off powerful gusts of wind. They can zip around while laden with more than half their weight in pollen or nectar. And altitude is no issue.

Bees, together with hummingbirds and fruit flies, produce great lift, achieve high power output and exhibit exquisite control, says Robert Dudley, an integrative biologist and head of the Animal Flight Laboratory at the University of California, Berkeley. "These are amazing design solutions to specific problems that engineers are super-interested in." As some scientists work to tease out the details of bee aerodynamics, others are applying the lessons learned to aeroplanes and even to create tiny flying robot bees with a host of potential applications in disaster response, espionage and agriculture.

It is a popular misconception that science cannot explain how bees fly. This conundrum can be traced back at least to 1934, when two Frenchmen, zoologist Antoine Magnan and his colleague André Sainte-Lague, did some calculations and concluded that it was impossible for bees to fly, despite the clear evidence

that they do. The problem lay not in the scientists' mathematics, but in their assumption that bees operate on the same principles as aeroplanes and gliders. "If you applied fixed-wing aerodynamic theory to the flight of a bee, you would have seen that the aerodynamics didn't work out," says Douglas Altshuler, a zoologist at the University of British Columbia in Vancouver, Canada. Aeroplanes rely on a steady flow of air, with the air above the wings moving faster than the air below, which generates lift.

DESCRIBING THE IMPOSSIBLE

Bees, of course, do not fly like aeroplanes — or even like most birds, which flap their wings up and down slowly. Bees beat their wings up to 240 times a second¹, which generates their noisy buzz and creates unsteady effects such as whirls and eddies in the air that surrounds them. If a plane created such a turbulent airflow, it would have problems, says Michael Dickinson, a zoologist and bioengineer at the California Institute of Technology in Pasadena. But in bees, such disturbances aid their ability to fly.

Some of these turbulent effects come from the angle of the wing, says Dickinson. Aeroplane wings are almost horizontal, generally deviating by less than 5°. At higher angles, wings create a leading-edge vortex — a tiny tornado turned on

its side — that initially provides an enormous amount of lift. But at the trailing edge of the wing, the air stream above fails to reunite with the one below and the effect disappears, causing loss of lift known as a stall.

Bees' wings hit the air at ever-changing angles, often greater than 50°, which provides the insects with high-lifting forces, Dickinson explains. At the end of each stroke, the wing moves in the opposite direction, generating a new, opposing leading-edge vortex that gives it another burst of lift. And because it takes longer for the streams of air to separate than for the wing to finish each stroke, the bee's flight does not stall (see "The secret to lift").

That mechanism is not the only source of lift, says Altshuler. At the end of the stroke, the bee actually flips its wing over, giving it a small amount of rotational lift too — a similar effect as putting backspin on a tennis ball. And because the wing has reversed direction, it is now travelling through disturbances from its previous stroke, so the air moves even more rapidly, enhancing the lift. "That is the wing recapturing its own wake," Altshuler says.

In addition to leading-edge vortices, rotational lift and wake recapture, bees can fine-tune their flight by varying the stroke length and, to a lesser extent, the speed². "It's almost like there's a menu of different aerodynamic

mechanisms,” Altshuler says, “and depending on how they position the wing stroke, they can select from this menu.”

HIGH FLIERS

Although the flight mechanics of bees are no longer mysterious, some of their aeronautical capabilities remain difficult to rationalize. Dudley travelled to Sichuan, China, to study honeybees³ living in the Himalayas at elevations of 3,250 metres. He placed the bees in a portable pressure chamber and used an air pump to remove some of the air and simulate higher altitudes. The bees lengthened their wing strokes and all were able to hover at pressures equivalent to 7,400 metres above sea level. The champion bee reached 9,125 metres — well above the peak of Mount Everest, at 8,848 metres. Dudley found that the same bees could fly at the equivalent of 1,000 metres below sea level. He does not have a good explanation for this range of ability. “You just have to wonder, why would this come up in nature?”

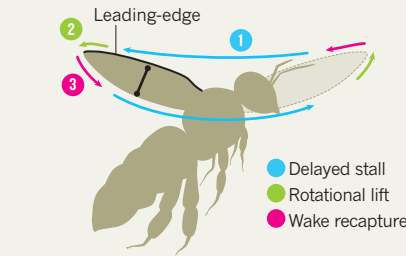
With such a suite of aerodynamic tricks, it is no wonder that engineers are using bees as inspiration in the design of aircraft. One thing that bees are very good at is coping with gusts of air carrying a force comparable to the speed and lift they are producing. The turbulence that a passenger jet experiences is many times smaller than its speed and lift, yet it causes a lot of discomfort to those on board; microbursts of air have caused planes to crash. A better understanding of how bees handle these forces might lead to new ways to cope with them in aeroplanes. “It’s teaching us how to harness some of the unsteady effects and flow that we try to squash out,” says Sean Humbert, head of the Autonomous Vehicle Laboratory at the University of Maryland in College Park.

A bee relies on tiny hairs covering its body to help it cope with different aerodynamic forces. Replacing existing aeroplane airflow sensors with ones that can detect localized forces, as a bee does, might give the plane’s systems more delicate control. If wind started to push the left wing upward, local sensors would detect those forces before they affected the rest of the plane, and the pilot could adjust the wing flaps so the plane would not rock. That sort of control will almost certainly be necessary, Humbert says, if a company ever wants to fly drones, or unmanned aerial vehicles, through the turbulent air of urban environments to deliver packages without crashing into roofs, playgrounds or other drones.

By carrying out experiments with bees in a wind tunnel, Humbert is learning how the insects’ sensors alter how they fly. Based on these insights, he is working with the aerospace industry to alter the designs of drones. Improved drones for urban use could come within a year or two, he says. Systems to smooth out passenger plane flights, which would need new sensors and motors that can react quickly enough, may take longer.

THE SECRET TO LIFT

Bees stay aloft through three main mechanisms. The process generates a burst of force during each phase.



1. Delayed stall

During each stroke, the leading edge of the wing creates vortices that remain attached to the wing and stop the bee from stalling.



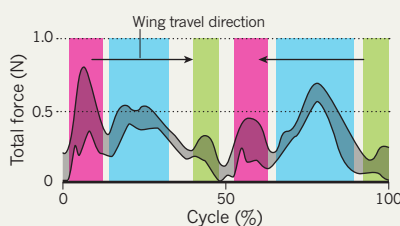
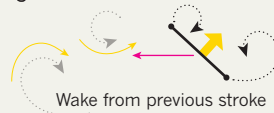
2. Rotational lift

But the wing also rotates between strokes to create lift in a similar — but not identical — way as a backspinning tennis ball.



3. Wake recapture

Finally, the bee’s wing is oriented on the upstroke such that it capitalizes on the wake of the preceding backstroke.



Some scientists are trying to build machines that fly in the same way as bees. “We’re trying to use the motions insects use, expecting to achieve similar forces and the ability to fly,” says Jeffrey Pulkamp, a mechanical engineer at the Army Research Laboratory in Adelphi, Maryland. A fleet of tiny, expendable airborne robots would have both military and civilian applications. They could, for example, fly into buildings that have collapsed following an earthquake, hurricane or bomb blast. “Imagine being able to send a team of autonomous little vehicles in there to sense people or temperatures or chemicals,” says Humbert, whose group is one of several working with Pulkamp’s team. Robotic bees could fit into spaces that larger drones cannot go,

slipping into crevices (and attracting less attention). Robotic bees could even be used as artificial pollinators, temporarily substituting for real bees — although because researchers are still in the early stages of creating these machines, that application could be 20 years away, says Robert Wood, head of the Robobees project at Harvard University in Cambridge, Massachusetts.

The first challenge was to get the tiny robots airborne. Pulkamp’s team has built prototypes. The imitation insects have bee-sized wings — and motors to flap them — made of a thin

“It’s teaching us how to harness some of the unsteady effects and flow that we try to squash out.”

film of the lightweight ceramic material lead zirconium titanate. The wings jut out from a thin sheet (the ‘body’) that is connected to a silicon chip by a tether that supplies power and keeps the platform

stable. With this set-up, the researchers have achieved hovering flight.

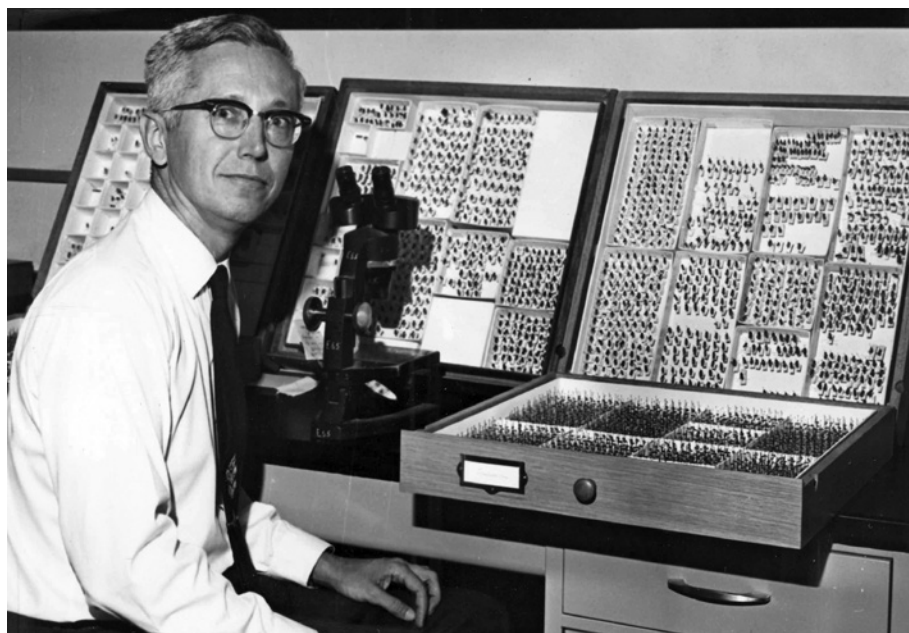
Wood’s Robobees project has also flown a tethered platform that has insect-scale wings made of a polymer membrane covered with spars of carbon fibre that resemble a bee’s wings more closely. Wood made the wings by copying the insect originals and not worrying too much about how they worked; they were more for testing his motor.

Further into the future, micro air vehicles may need capabilities such as chemical sensors, infrared detectors and microscopic versions of Geiger counters to carry out tasks. Because the micro vehicles will be untethered, they will also probably require radio transmitters to communicate and global positioning system equipment. And to run the whole package, they will need computer control, all of which points to one of the trickiest aspects of making tiny autonomous flying robots: power. Five or six years ago, Wood says, a good polymer lithium battery that was small enough to fit on to a robobee provided enough power for 19 seconds of flight. Advances in the propulsion system, energy storage, motors and electronics means that Wood’s estimate has increased to a few minutes. “But it still highlights the challenges with power for these small robots,” he says.

All of these challenges — flight, navigation, control, power — will require several more years of work if scientists are to replicate the abilities that natural selection bestowed on bees. It took 50 years to explain how bees fly, and scientists still have not discovered all their secrets. ■

Neil Savage is a freelance science and technology writer in Lowell, Massachusetts.

- Altshuler, D. L., Dickson, W. B., Vance, J. T., Roberts, S. P. & Dickinson, M. H. *Proc. Natl Acad. Sci. USA* **102**, 18213–18218 (2005).
- Vance, J. T., Altshuler, D. L., Dickson, W. B., Dickinson, M. H. & Roberts, S. P. *Physiol. Biochem. Zool.* **87**, 870–881 (2014).
- Dillon, M. E. & Dudley, R. *Biol. Lett.* <http://dx.doi.org/10.1098/rsbl.2013.0922> (2014).



Q&A Charles Michener

A life with bees

Charles Michener has been studying bees for more than 80 years, and, although he has seen many changes in the field, his interest in these insects has not diminished. Now aged 96, he contributes to bee research as a Watkins distinguished professor emeritus at Kansas University in Lawrence.

What sparked your interest in bees?

I was brought up in Pasadena in southern California, where I spent hours studying wild-life. I drew as many of the native flowering plants in bloom as I could. At about age ten, I ran out of new plants to draw, so I started collecting and drawing insects. I made more than 1,200 sketches, many of them of various types of bee. There was one bee in particular that probably led to my fixation. It was a minute *Perdita rhois* Cockerell, a beautiful yellow-and-black insect that was attracted to the daisies in our yard every summer.

How did you turn your childhood fascination into a research career?

After completing my PhD in North American bee genera at the University of California, Berkeley, I spent several years studying other creatures. In 1942, I was hired by the American Museum of Natural History in New York to work on Lepidoptera — butterflies and moths. However, my supervisor liked the fact that I had my own interest, and so allowed me to continue researching bees and to publish my findings. He had previously worked on bees, which probably played some part in allowing me to continue, but I suspect that even had I been working on beetles, he would have welcomed my study.

In 1943, I volunteered for the Army Sanitary Corps. My first assignment was in Mississippi on mosquito-borne diseases. After a year I was sent to Panama to study the chigger mites (Trombiculidae) that were transmitting scrub typhus in the Pacific and hurting the US war effort by taking people out of action. As I travelled, I continued to collect bees and even publish papers — including one of the first papers on bees in Mississippi. While I was in Panama, I saw stingless and orchid bees everywhere — it was my first experience with neotropical bees and led to the publication in 1954 of *Bees of Panama* (Literary Licensing).

When you left the service, what attracted you to Kansas?

I realized that to undertake bee biology studies, I needed experience in wild-bee behaviour and nesting habits. The University of Kansas offered this opportunity, so I moved to the state in 1948 and have never looked back. My work there enabled me to contribute to finding solutions for pollination problems, especially in alfalfa seed, which has a particular connection with bees. Without bees, alfalfa produces little seed. In the presence of honeybees, alfalfa produces some seed — but nowhere near optimum. The maximum yield is produced when native bees — bumblebees and leaf-cutter bees — pollinate

the plants. Our solution was to provide nesting materials in the fields for the native bees. For bumblebees, this meant small boxes in which they could establish colonies. This is now the standard method for farming alfalfa seed.

What are the biggest changes you have seen in bee populations and in bee research?

In earlier years, people rarely collected numerical data on populations. They noted when a bee species was abundant or scarce, and that was about it. So to show any change over the years is difficult because initial data are not there.

People are still discovering new species of bee. In some cases this may be the result of a species simply migrating to a new area, but usually there is no way to know. Bees are highly susceptible to changes in the natural environment. The geographical range of various species will have to be tracked more carefully, which will be difficult to do.

In terms of research, there have been a host of changes within studies of social behaviour and understanding of caste determination, that is, among workers, drones and queens. But the biggest change is the ability to relate everything to systematics — the organization of bees into families, genera and species based on morphology and now on DNA. For example, honeybees nest in large families; leaf-cutter bees nest alone in burrows in the ground or in wood. Every feature of bee behaviour, physiology and ecology helps to reinforce the classification based on morphology.

How has the perception of bees changed throughout your career?

People have a better understanding now of the economic value of bees, because researchers like me are talking about the importance of bees for crops. And it is not just about the honeybee, of which there are only nine species. There are hundreds of other bees that have an impact through pollination on agriculture and natural vegetation, which were previously under-appreciated.

What are the main challenges faced by bees?

The biggest problems are ecological, such as the destruction of natural vegetation. More and more studies indicate that human interferences in the environment — such as through climate change and spraying pesticides on flowering crops — are strongly influencing the decline in abundance or the extermination of various bees.

It has been 80 years since the publication of your first paper. Will there be any more?

Yes. I have a paper in press, due to be published this year, about a new species of bee in Thailand discovered by my collaborator and former graduate student, Natapot Warrit. ■

INTERVIEW BY JULIE GOULD